Application of Pattern Recognition Techniques for Early Warning Radar (EWR) Discrimination

Progress Report II

27 January 1995

Mr. Thomas Foster
Principal Investigator

SenCom CORPORATION

Contract No. N00014-94-C-0240
SBIR BMDO 94-003
CDRL No. 0001AA

SenCom Corporation
3 Preston Court
Bedford, MA 01730

Research supported by Ballistic Missile Defense/Innovative Science and Technology and managed by the Office of Naval Research

OFFICE OF THE UNDER SECRETARY OF DEFENSE (ACQUISITION)
DEFENSE TECHNICAL INFORMATION CENTER
CAMERON STATION
ALEXANDRIA, VIRGINIA 22304-6145

Feb 27, 1995

IN REPLY REFER TO DTIC-OCC

SUBJECT: Distribution Statements on Technical Documents

TO: Office of the Chief of Naval Research
800 north Quincy Street
Arlington, VA 22217-5000
Code 22 ~~AL Dumas~~

1995 0925 071

1. Reference: DoD Directive 5230.24, Distribution Statements on Technical Documents, 18 Mar 87.

2. The Defense Technical Information Center received the enclosed report (referenced below) which is not marked in accordance with the above reference.

N00014-94-C-0240
Application of Pattern Recognition Techniques for Early Warning Radar (EWR) Discrimination
Jan 27, 1995

3. We request the appropriate distribution statement be assigned and the report returned to DTIC within 5 working days.

4. Approved distribution statements are listed on the reverse of this letter. If you have any questions regarding these statements, call DTIC's Cataloging Branch, (703) 274-6837.

FOR THE ADMINISTRATOR:

1 Encl

GOPALAKRISHNAN NAIR
Chief, Cataloging Branch

FL-171
Jul 93

DISTRIBUTION STATEMENT A:

APPROVED FOR PUBLIC RELEASE: DISTRIBUTION IS UNLIMITED

DISTRIBUTION STATEMENT B:

DISTRIBUTION AUTHORIZED TO U.S. GOVERNMENT AGENCIES ONLY; (Indicate Reason and Date Below). OTHER REQUESTS FOR THIS DOCUMENT SHALL BE REFERRED TO (Indicate Controlling DoD Office Below).

DISTRIBUTION STATEMENT C:

DISTRIBUTION AUTHORIZED TO U.S. GOVERNMENT AGENCIES AND THEIR CONTRACTORS; (Indicate Reason and Date Below). OTHER REQUESTS FOR THIS DOCUMENT SHALL BE REFERRED TO (Indicate Controlling DoD Office Below).

DISTRIBUTION STATEMENT D:

DISTRIBUTION AUTHORIZED TO DOD AND U.S. DOD CONTRACTORS ONLY; (Indicate Reason and Date Below). OTHER REQUESTS SHALL BE REFERRED TO (Indicate Controlling DoD Office Below).

DISTRIBUTION STATEMENT E:

DISTRIBUTION AUTHORIZED TO DOD COMPONENTS ONLY; (Indicate Reason and Date Below). OTHER REQUESTS SHALL BE REFERRED TO (Indicate Controlling DoD Office Below).

DISTRIBUTION STATEMENT F:

FURTHER DISSEMINATION ONLY AS DIRECTED BY (Indicate Controlling DoD Office and Date Below) or HIGHER DOD AUTHORITY.

DISTRIBUTION STATEMENT X:

DISTRIBUTION AUTHORIZED TO U.S. GOVERNMENT AGENCIES AND PRIVATE INDIVIDUALS OR ENTERPRISES ELIGIBLE TO OBTAIN EXPORT-CONTROLLED TECHNICAL DATA IN ACCORDANCE WITH DOD DIRECTIVE 5230.25, WITHHOLDING OF UNCLASSIFIED TECHNICAL DATA FROM PUBLIC DISCLOSURE, 6 Nov 1984 (Indicate date of determination). CONTROLLING DOD OFFICE IS (Indicate Controlling DoD Office).

The cited documents has been reviewed by competent authority and the following distribution statement is hereby authorized.

A
(Statement)

OFFICE OF NAVAL RESEARCH
CORPORATE PROGRAMS DIVISION
ONR 353
800 NORTH QUINCY STREET
ARLINGTON, VA 22217-5660

(Controlling DoD Office Name)

(Reason)

(Controlling DoD Office Address, City, State, Zip)

DEBRA T. HUGHES
DEPUTY DIRECTOR
CORPORATE PROGRAMS OFFICE

(Signature & Typed Name)

(Assigning Office)

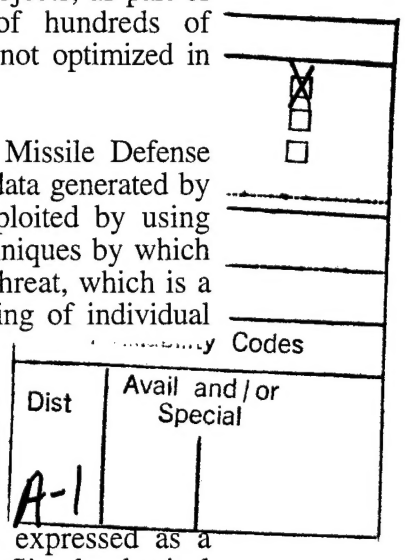
15 SEP 1995

(Date Statement Assigned)

Introductory Remarks

The PAVE PAWS and the Ballistic Missile Early Warning System (BMEWS) radars are phased array radars which support the Early Warning System (EWS). These radars operate at UHF and were designed to detect and track large numbers of objects, as part of the then perceived threat, i.e., a massive ballistic missile attack of hundreds of ICBMs/SLBMs. Discrimination and tracking of individual objects were not optimized in their design.

In our recent work over the past two years in support of the Ballistic Missile Defense Organization (BMDO), we have noted that the radar cross section (RCS) data generated by a PAVE PAWS radar appears to contain information that might be exploited by using pattern recognition techniques. It is our purpose to study and present techniques by which this information can be extracted and exploited against today's perceived threat, which is a limited ballistic missile attack, during which the discrimination and tracking of individual objects becomes a prime requirement.



Background

The basic idea we wish to exploit is that when the RCS amplitude is expressed as a function of time, we see a global structure that can be viewed as a pattern. Simple physical considerations suggest that different classes of objects will produce different patterns in the RCS. For example tanks (first or second stages) are large objects and generally have a tumbling motion associated with their trajectory. This produces patterns that exhibit large means and variances. Just the opposite is true for reentry vehicles (RVs). In this case we are considering small objects which tend to be spin stabilized. This implies patterns with small means and variances. While such differences might not be apparent within a small sampling of data points, such patterns should and do manifest themselves over a period of hundreds of seconds.

Of course the reality of the situation is more complicated. We also need to consider the problem of RVs which may be executing an effective tumbling motion, at least relative to the radar line of sight. In addition we must also realize that there may be other objects that are approximately the same size as the RV and that a more detailed look at the RCS pattern may be necessary to distinguish these cases.

This relates back to a technical issue that was discussed in Progress Report I (PR I). The wavelength at which these radars operate is comparable in length to an RV or other similarly sized object. This means that some small scale structure which would distinguish the RV from the nose fairing, for example, cannot be effectively sampled by radars operating at UHF. The question then becomes how to distinguish between various classes of electrically small objects. There appears to be two ways in which that might be done. First, one can postulate that over a long period of time, perhaps several hundred seconds, the changing aspect or viewing angle will induce visible changes in the RCS pattern that will distinguish the RV from other small objects. Secondly, it may be that while the operating wave length is relatively insensitive to the precise shape of the illuminated object, the various RCS patterns will be different in small details which can be examined by considering the higher moments of the distribution of RCS returns.

Objective

As stated in PR I, our fundamental task is to assess the feasibility and application of simple pattern recognition algorithms and techniques to the problem of radar target discrimination and classification at UHF frequencies. The aim is to develop a low cost and low risk method to improve EWR discrimination. The initial step was to explore ways in which to present the data prior to applying pattern recognition techniques. The principal results of this work were discussed in PR I. The immediate task is to examine more data in order to accomplish two objectives: (1) assess the utility of various data representations; and (2) begin a division of the data base into a small set of classes based on results from the various data representations. This work has started and some preliminary results are presented in this report.

There are two major parts to this report, a section on our initial classification procedure and an appendix in which we present a RCS modeling technique which offers a possible way to analyze a tumbling RV. In the initial classification section, we review the approach used to begin a systematic examination of our PAVE PAWS RCS data base. We then discuss a methodology that could be used to extract and identify the various patterns present in the RCS data base. Some preliminary findings are presented. In the Appendix A, we present an approach for modeling the RCS returns from simple shaped objects undergoing a specified motion. We also present a possible pattern recognition technique based on the RCS modeling.

Initial Classification

The work presented in PR I considered the results of applying a number of data simplification schemes to a limited sampling of data sets. These schemes or data representations were discussed in the first report, so only a brief summary of them is repeated here.

Cumulative probability

A first approach is to consider the entire data set and construct histograms and cumulative probability distributions (CPDs) for each set. The cumulative distribution is formed by calculating the percentage of RCS amplitudes that are greater than any particular amplitude value within the given range of the data set. The histogram is simply the result of dividing the amplitude range into bins and counting the number of data points that fall into each bin. The shape of the distribution or histogram depends on the character of the original data set.

Fourier coefficients

A second approach also considers the data set as a whole. In this case one generates the real and imaginary coefficients of the Fourier transformation of the RCS time histories. The values of the coefficients are then plotted in the complex plane. The relative spread of the points becomes a measure of the amplitude strength as a function of the frequency content. One can also consider the spectral content of the data in more detail.

Number density

Third, we consider a number density approach. In this case we use a reduction process where the amplitude is not the key discriminant. The amplitude time space of an arbitrary

RCS time plot is divided into a fixed number cells, and the number of data points falling within each cell are counted. The cells and their counts can then be represented as a matrix.

Various pictorial analyses are possible. These can take the form of either two dimensional surface maps or three dimensional column plots. What is seen in these cases is the degree of structure present in the data. This suggests the possible use of entropy methods to do pattern recognition.

N point reduction

As a final approach, the original data set is divided into subsets of N points each. Here N can be any suitable integer. However for the analysis thus far, N has been set to twenty. Then for each subset, the mean and standard deviation are calculated. It is found that this approach illuminates the general character of the original data set in terms of averages and variability in a fairly simple and effective way.

These techniques are summarized below in Table 1.

Technique	Variables	Discriminant
Cumulative Probability Distribution	Percent & Amplitude	Shape of Distribution
Fourier Transform	Real & Imaginary Coefficients	Radius Length
Number Density	Number of Points per Cell	Shape & Structure of Distribution
N Point Reduction	Mean & Standard Deviation	Relative Position & Pattern

Table 1. Summary of Processing Techniques

Data Format and Selection

Our initial results in PR I were based on the examination of a limited number of data sets. The next step was to enlarge the data base under consideration. The initial results suggested that there were a small number of basic patterns in the data and that the various data processing techniques allowed for a clear illustration of these. As more data has been processed, this has remained true, but additional subtleties and complications have also become evident. Before presenting our new findings, we will first review the data analysis procedure.

Conceptually one can divide the RCS data base into three groups (i.e., groups I, II and III) corresponding to the date on which they were collected by the radar. Each set has the following information: the track file ID number; the time; the X, Y and Z position of the object in radar face coordinates; and the RCS. Thus the position and RCS are both given as functions of time. Furthermore, the range, azimuth and elevation of the object are also easily obtained through simple transformations involving the radar face coordinates.

At this stage, we have made an initial survey through groups I and II. Precisely what is meant by this will be made clear shortly. All of the data sets within group I have been visually examined, amounting to nearly 60 sets. However, a significant fraction of these are either deficient in the number of sample points or do not span a sufficient length of time, i.e., 100 seconds or more, to be consistent with the global nature of our approach.

As a result only 28 data sets from group I were appropriate for detailed analysis. For group II, we were more selective right from the start, in that we generally demanded a minimum track data rate of 0.25 Hz and a period of observation of 100 seconds or so before we would even visually examine the data set[1]. With these selection criteria the number of processed data sets from group II totaled 24. In addition, we anticipate that approximately 40 or more data sets from group III will be processed during the next several weeks.

Description of MATLAB Environment

As a tool to generate and assess various data representations, we are using MATLAB, a numeric computation and visualiztion software package. There are a number of advantages to using MATLAB. First of all it presents a natural way of emulating the entire pattern recognition system. It has a powerful Graphical User Interface (GUI) capability that allows for the generation of menu driven processes and various types of controls (push buttons, sliders, etc.) which ultimately control an excellent graphics package. Thus the entire pattern recognition process can be simulated within MATLAB. This simulation can be run in an automatic mode or with an observer in the loop, since all intermediate data and displays can be brought to the screen. This is of course very useful during the algorithm development and debugging phases.

Furthermore, any additional FORTRAN code to be developed, including that which has been developed for the RCS modeling and simulations (discussed in Appendix A) can be directly incorporated into the MATLAB environment as MEX (executable) files. These can then be called within the MATLAB simulation. Incorporating the data processing routines developed earlier represent a first step in this direction.

On a more practical and immediate level, MATLAB simply offers an easy way of surveying a large amount of data. Data sets from each group are loaded into corresponding "work spaces" in the MATLAB environment. Each file is then run against a script which works through the various data processing techniques (plus additional related plots) and outputs results to the screen. The following figures (1 through 7) are examples to illustrate the utility of MATLAB and to show typical outputs.

[1] In a few cases we accepted data sets with an observation time of slightly less than 100 seconds if the track rate was significantly above 1 Hz.

Figure 1 presents the basic picture of the data, i.e., the RCS amplitude as a function of time, given in seconds relative to the start time of the event.

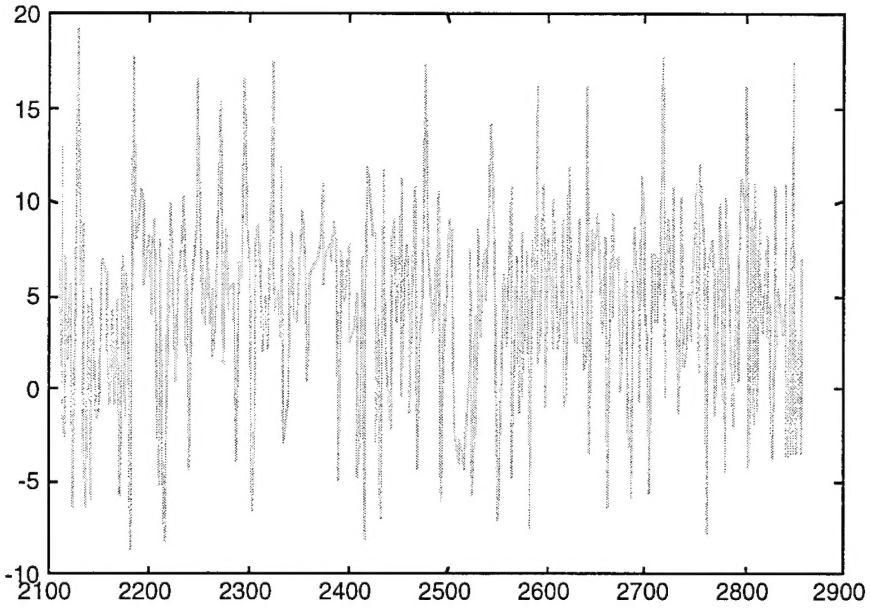


Figure 1. RCS Vs Time (Object 2348)

The time is given by the horizontal axis, while the dependent variable, in this case the RCS amplitude, is given by vertical axis.

The next step in the MATLAB script produces a histogram, a probability density function (PDF) and a cumulative probability distribution (CPD) of the data. The width of the bins of the histogram were set at 0.5 dB. The PDF is essentially a smoothed version of the histogram with the entry in each bin normalized by the total number of points. Thus the total area under the PDF curve is equal to unity. These plots are outputted together and shown in figure 2.

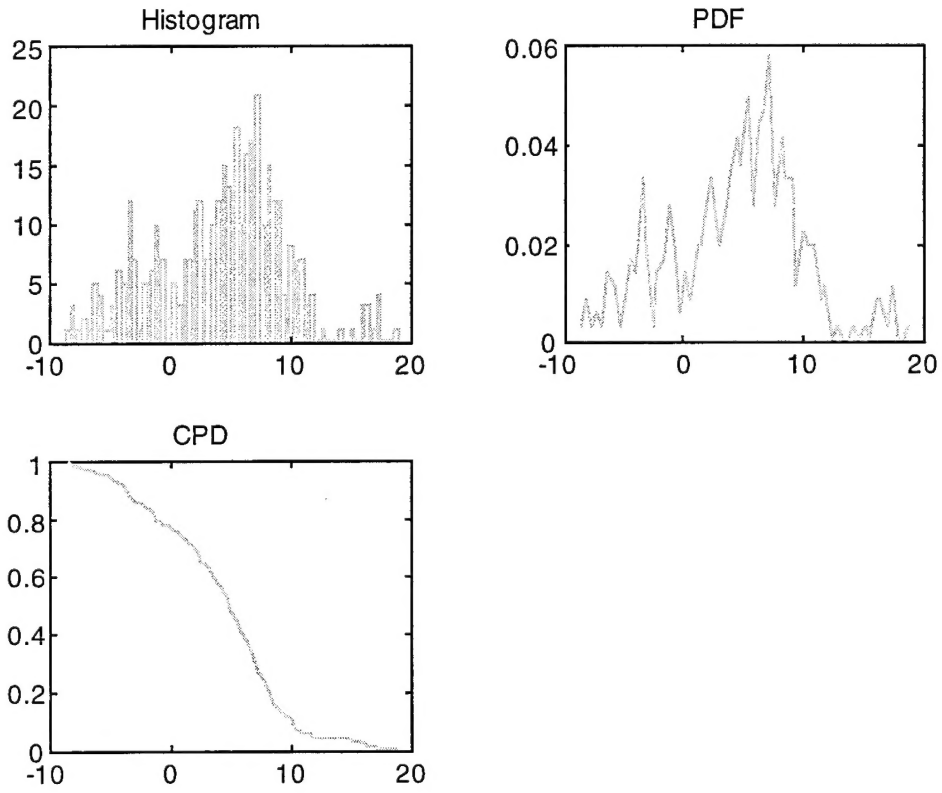


Figure 2. Histogram, PDF and CPD (Object 2348)

The Fourier coefficients (at two scales) are given below in figure 3.

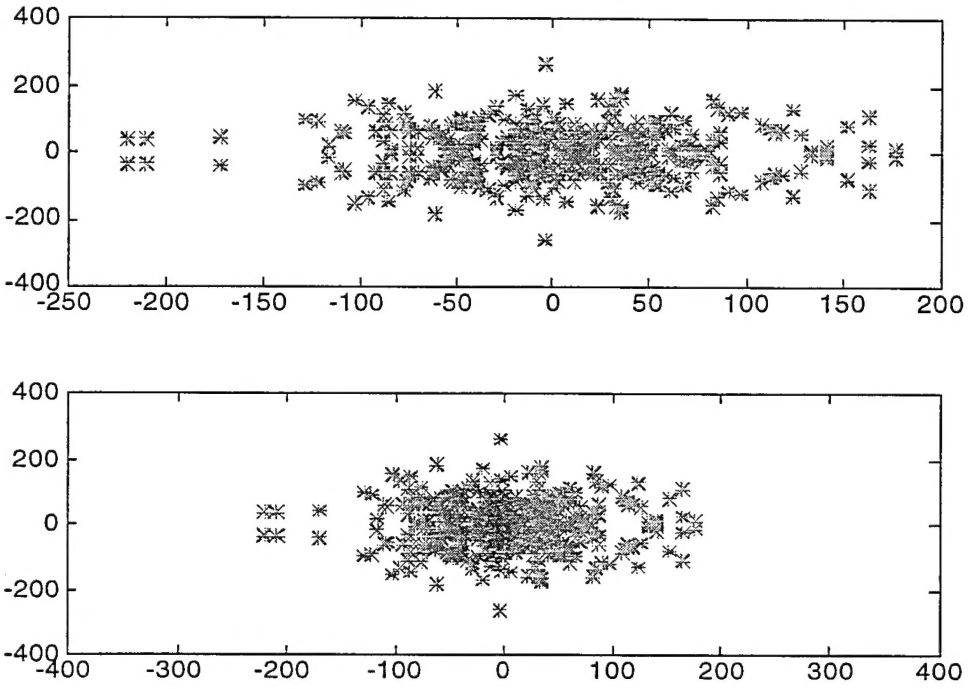


Figure 3. Fourier Coefficients (Object 2348)

In figure 4 we are given the power or the amplitude squared of the data (given by the vertical axis) as a function of both the frequency (in Hz) and the period (in seconds), given by the horizontal axes. For the latter, it is presented at two different scales.

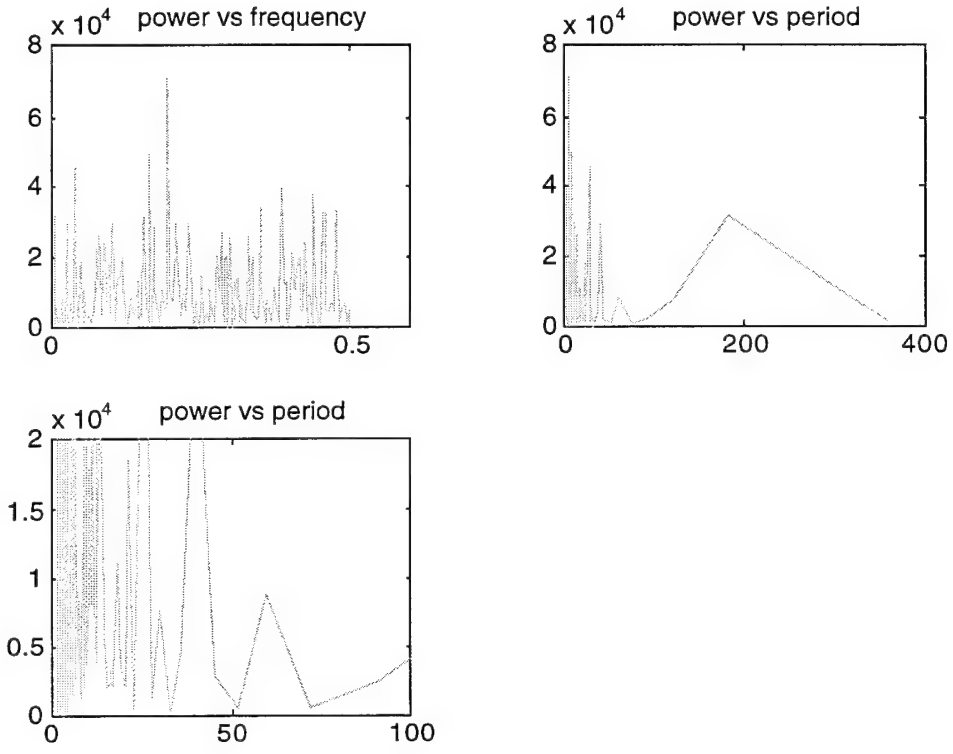


Figure 4. Spectral Data (Object 2348)

In figure 5 below, we show the distribution of the 20 point means and standard deviations. The mean is measured along the horizontal axis and the standard deviation is given along the vertical axis.

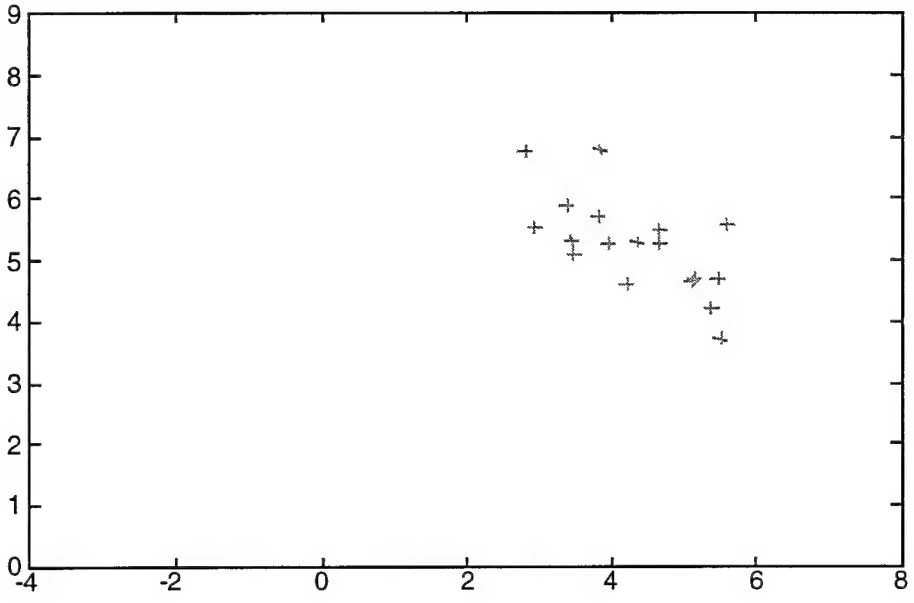


Figure 5. 20 Point Means Vs 20 Point Standard Deviations (Object 2348)

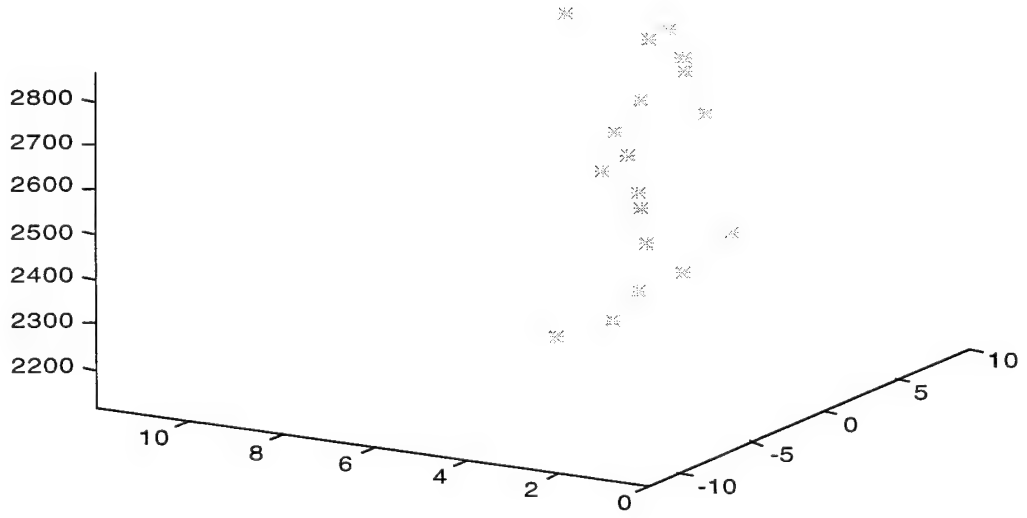


Figure 6. 20 Point Means and Standard Deviations as a Function of Time (Object 2348)

Figure 6 shows a three dimensional plot which gives the time variation of the 20 point means and standard deviations. This plot is actually the end result of an animation

sequence that is presented to the observer in the MATLAB environment. As such there are better ways of presenting this information and we will not work with this further.

Instead, as we see in figure 7, we the take two dimensional projections through this data and show the 20 point means and the 20 point standard deviations as explicit functions of time, where time is measured along the horizontal axes.

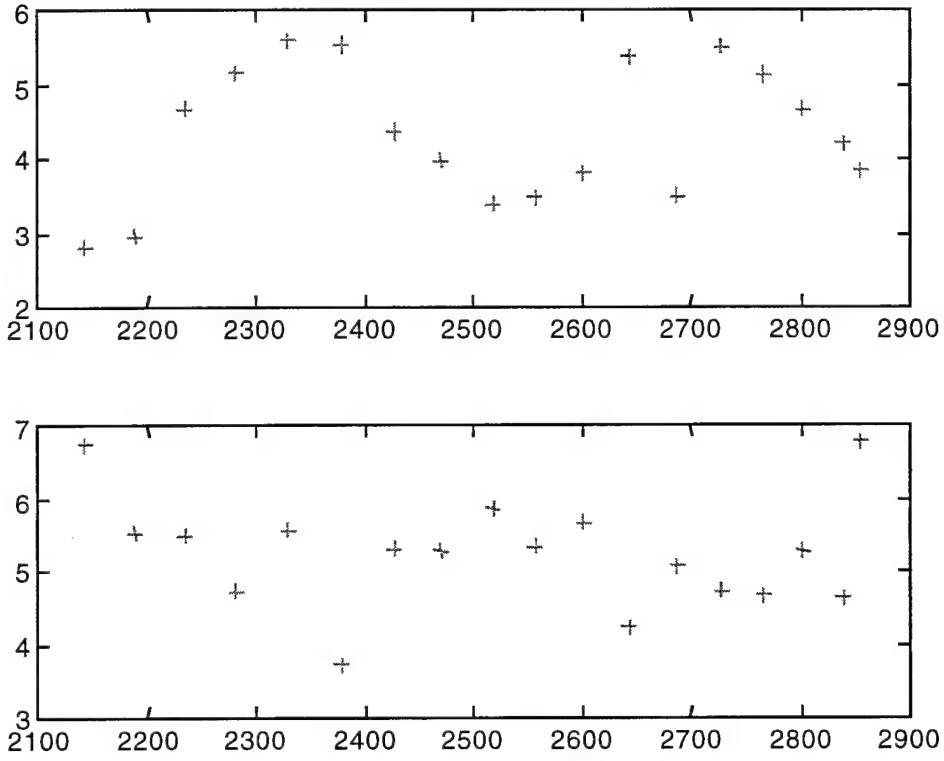


Figure 7. 20 Point Means and 20 Points Standard Deviations Vs Time (Object 2348)

Initial Data Processing

A sample of data sets from groups I and II have been processed through the MATLAB environment. Initially the data was passed through a primitive classifier based on the N point reduction scheme. That is, each data set was divided into subsets of 20 points and the mean and standard deviation of each subset were calculated. The resulting means and standard deviations were plotted. The plot space itself was then divided into four regions (classes 1, 2, 3 and 4) which, based on our understanding of the physics of the process,

corresponded roughly to four classes of objects: i.e., RVs, tanks, PBVs and fragments. The data set was then assigned to the class whose region contained most of the data points.

Some remarks are required. First, this is not the classifying scheme conceptualized in PR I. In that case the mean-standard deviation space would be divided into small cells. Then objects, whose classifications were known could be used to assigned the probability that a particular cell in that space was occupied by an object of a certain type (e.g., a tank or RV). However, in the present situation, we are simply dividing the space into large regions based on some obvious cases and intuition. However one could envision this as a way to make a preliminary cut at the classification problem.

The output from this method could be analyzed by other means and based on that, a final classification of the object could be made. This would be in the spirit of the classification architecture presented in the first progress report. Note also that it would allow at least one part of the classifier to be adaptive in the sense that it would not be necessary to know the precise region (or cells) within the mean-standard deviation space that particular object classes should occupy. One could simply start with a reasonable guess and let the results from the other processing schemes educate the N point processor.

However, our aim here is more modest. The first question we address is whether the other processing schemes are sufficiently robust to add anything beyond what is determined from the initial classification based on the subset means and standard deviations. That is, based on this approach, does everything initially classified as a RV, for example, also look like a RV within the other processing schemes. If it does, then there may be a problem because we will see that class 1 data includes data sets with different RCS patterns. If our additional techniques do not key on these differences then nothing is gained by considering them.

To summarize, the point is to divide the data (from groups I and II) into four sets or classes (1 through 4) and then hopefully divide the data into additional sub-classes or types based on runs through the MATLAB environment.

Class 1 Data

For this report, we will concentrate principally on the class 1 or RV objects, as we have not yet had time to carefully look at the data assigned to the other classes. For now, we suspect that the data assigned to class 2 (tank) will present few surprises. Class 1 data is data that resides primarily within region 1 of the mean-standard deviation space. Region 1 is the area defined by a mean between 0 and 6 and a standard deviation from 0 to 3.5. To be consistent with PR I this region would correspond to an RV, although we might suspect that other classes of objects could reside within that region under certain conditions. The question thus becomes whether we have the tools to detect the presence of these other classes within region 1.

To orient ourselves we start with the track file on object 2609 which was discussed briefly in PR I and tentatively identified as an RV. Figures 8 and 9 show the RCS vs. time plot and the 20 point means vs. 20 point standard deviations plot, respectively.

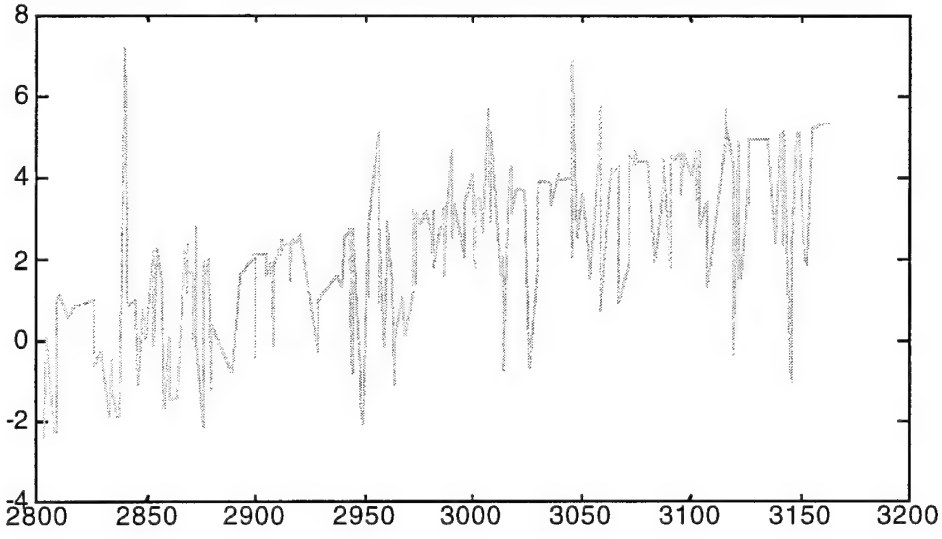


Figure 8. RCS Vs Time (Object 2609)

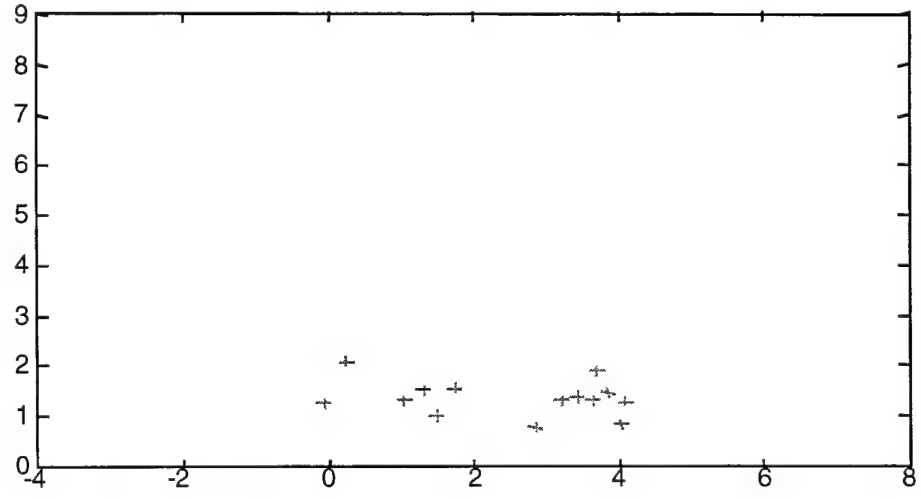


Figure 9. 20 Point Means Vs 20 Point Standard Deviations (Object 2609)

Figure 10 shows the histogram and related plots for the same object.

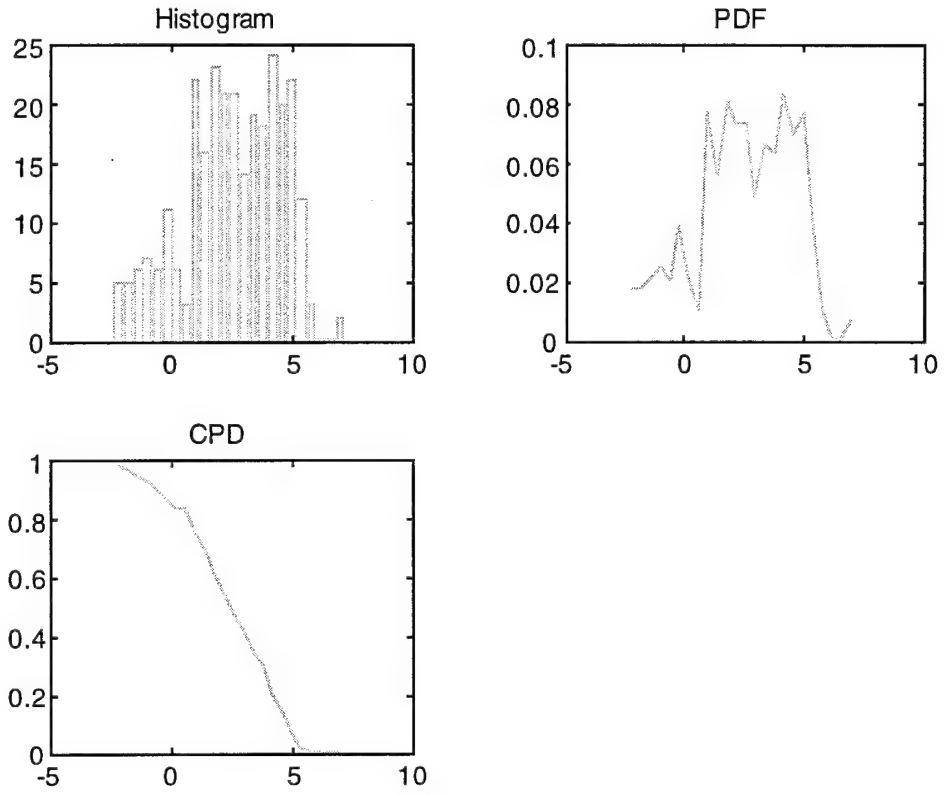


Figure 10. Histogram, PDF and CPD (Object 2609)

The histogram in figure 10 indicates a narrow variance with most of the values falling between 0 and 5 dB. Since the PDF is essentially a smoothed out version of the histogram, it displays the same basic shape as the histogram. However, it may present an easier pattern for an algorithm to identify. The CPD represents a more concise picture indicating data with a narrow variance and small amplitudes. However the information may be more readily apparent in the histogram and PDF. In particular, we see an asymmetry due to the large number entries into high amplitude bins. This is a reflection of the systematic increase in the RCS signature over time, which is due to a slow variation in the aspect angle as discussed in PR I.

Figure 11 shows the Fourier coefficients with two scales for the horizontal (real) axis.

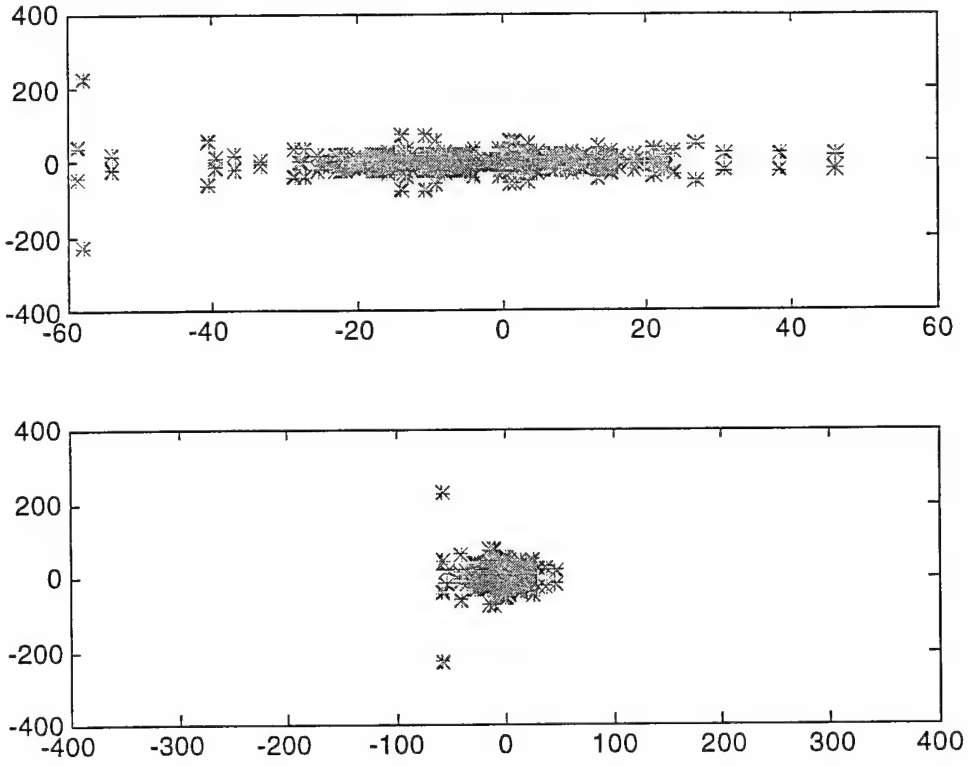


Figure 11. Fourier Coefficients (Object 2609)

What is apparent in this figure is that the real coefficients are tightly clustered and in fact with the exception of two points, the imaginary coefficients (vertical axis) are also tightly clustered. As discussed previously, this is a typical result given the above RCS time history.

Additional spectral information is presented in figure 12. Here we are considering the amplitude squared (power) as a function of the frequency and as a function of the period. In the later case the data is presented at two scales.

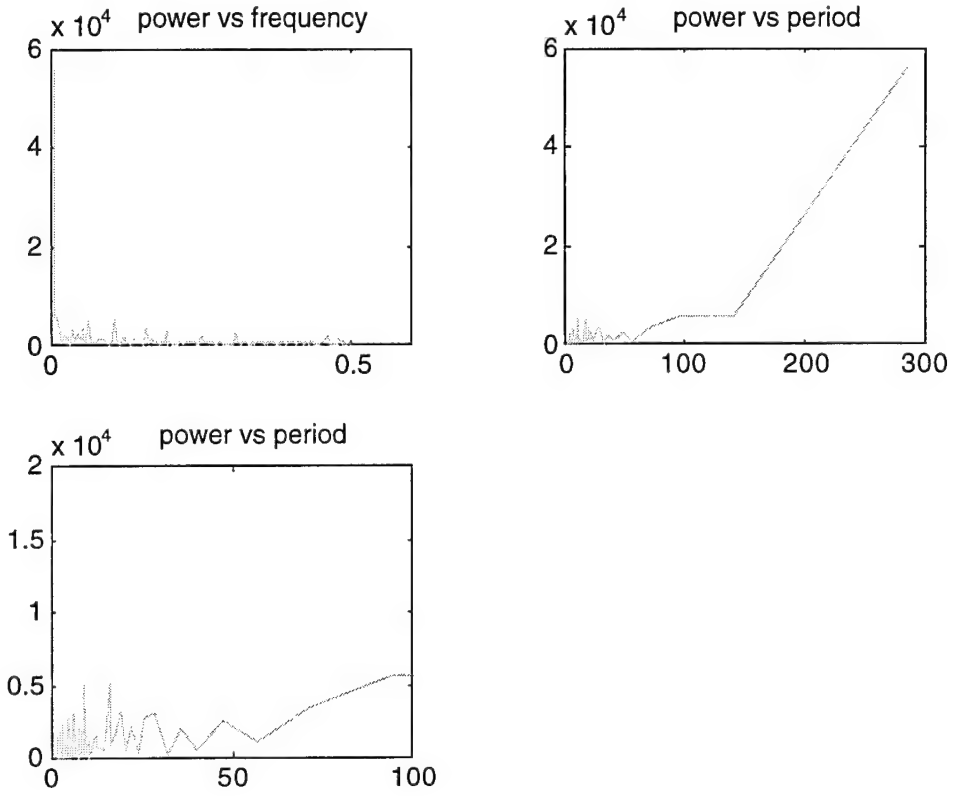


Figure 12. Power Vs Frequency and Period (Object 2609)

If we first consider the frequency response, we note that most of the power is taken up by a near zero Hz line. This is due to the systematic increase of the RCS over time or aspect angle. In fact one can almost draw a straight line through the data set. This straight line or 0 Hz characteristic is reflected also in the power vs. period graph which is the origin of the monotonically increasing behavior of the plot as the period becomes large. In this situation one begins to sample the global spectral character of the data.

For completeness, figure 13 shows the time variation of the means and standard deviations for this object. Here we see that the increase in the RCS is reflected in the increase of the 20 point means as a function of time. However the variance or standard deviation remains small and relatively constant over the data collection time.

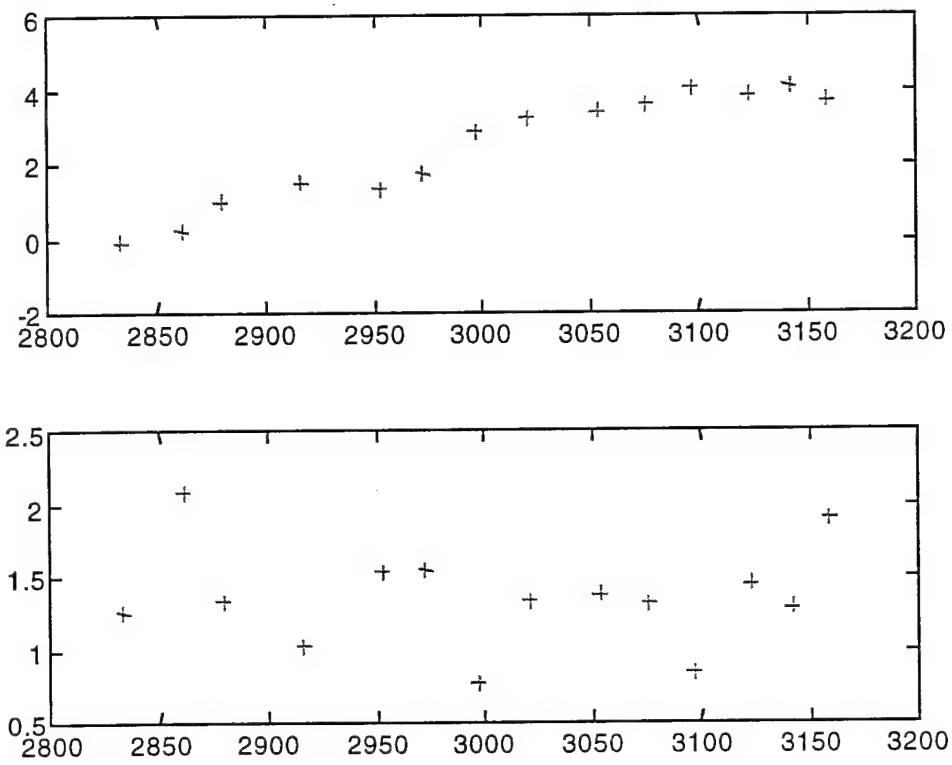


Figure 13. 20 Point Means and 20 Point Standard Deviations Vs Time (Object 2609)

During our initial pass through the two data groups, we have discovered a number of cases that have RCS patterns that are similar to the above object. Examples of these are shown in figures 14 and 15.

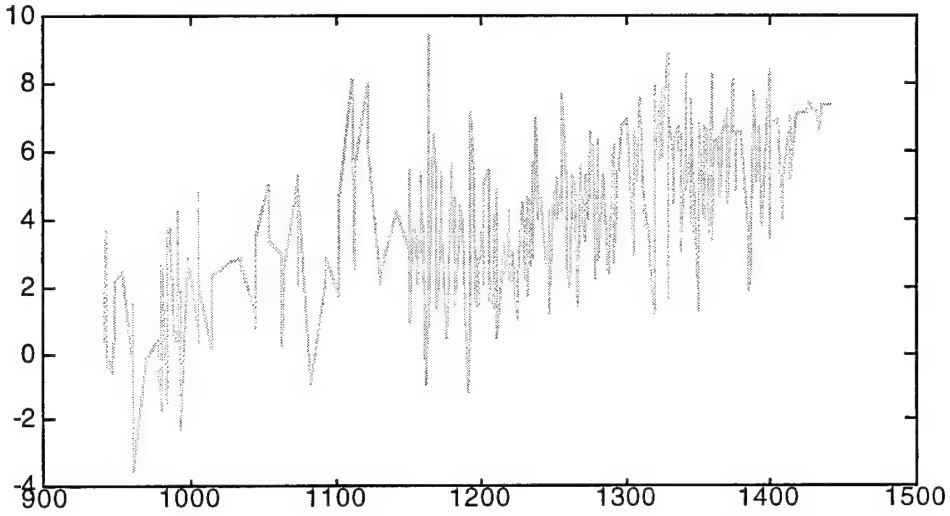


Figure 14. RCS Vs Time (Object 1221)

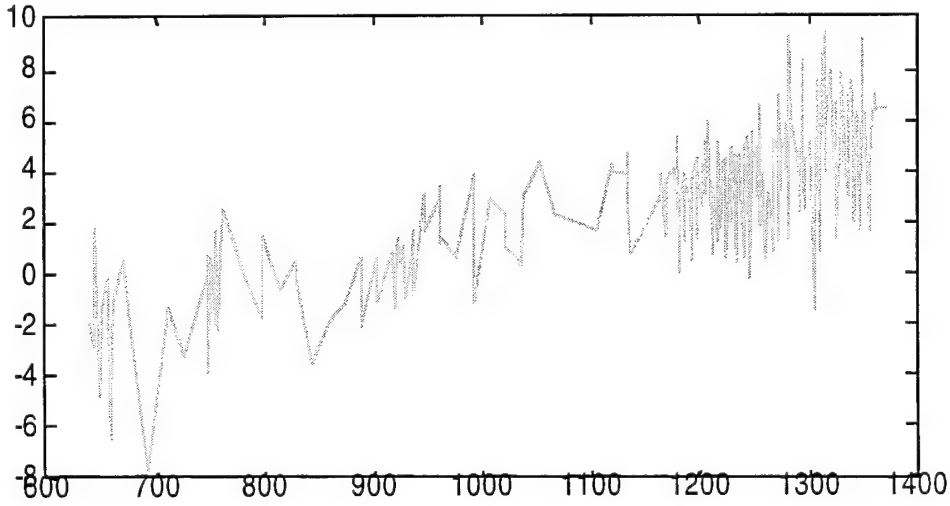


Figure 15. RCS Vs Time (Object 1193)

At this point, since the histogram and spectral plots for these two cases are quite similar to that seen for object 2609, it is more illuminating to consider an object that is classified as a class 1 object but whose RCS pattern appears different.

For this case, we consider object 1234 whose RCS time history is given below in figure 16.

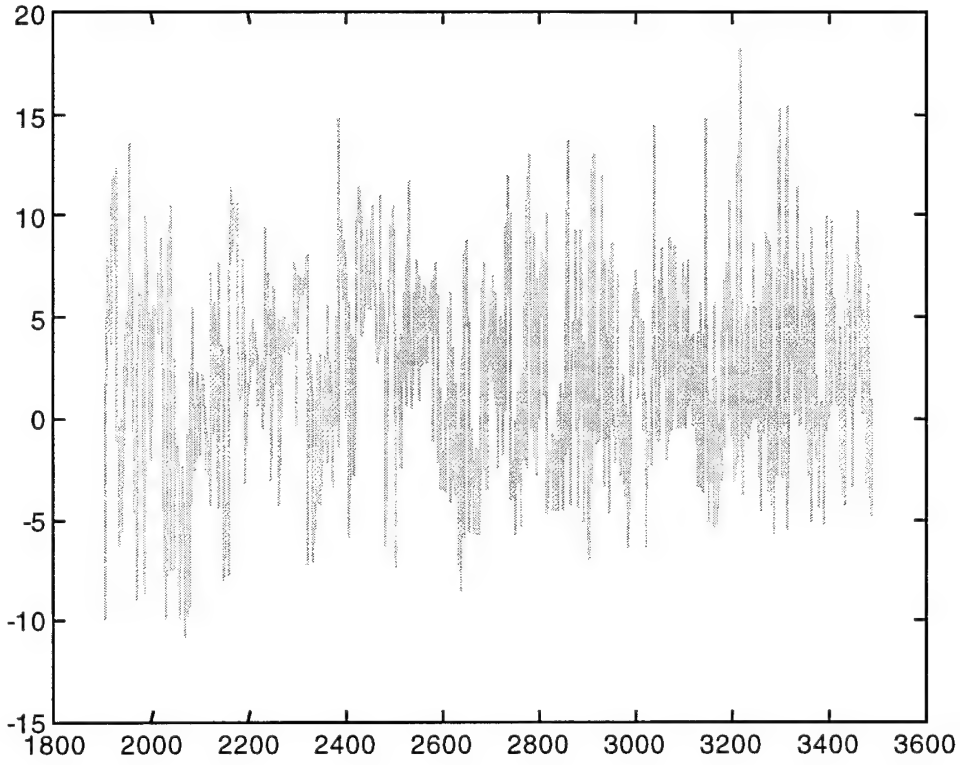


Figure 16. RCS Vs Time (Object 1234)

While our primitive classifier assigned this data set to class 1, it is obvious that the RCS signature is different from what we saw for object 2609 (see figure 8). The question is whether this difference is reflected in the histogram and spectral plots.

The histogram data is reproduced in figure 17.

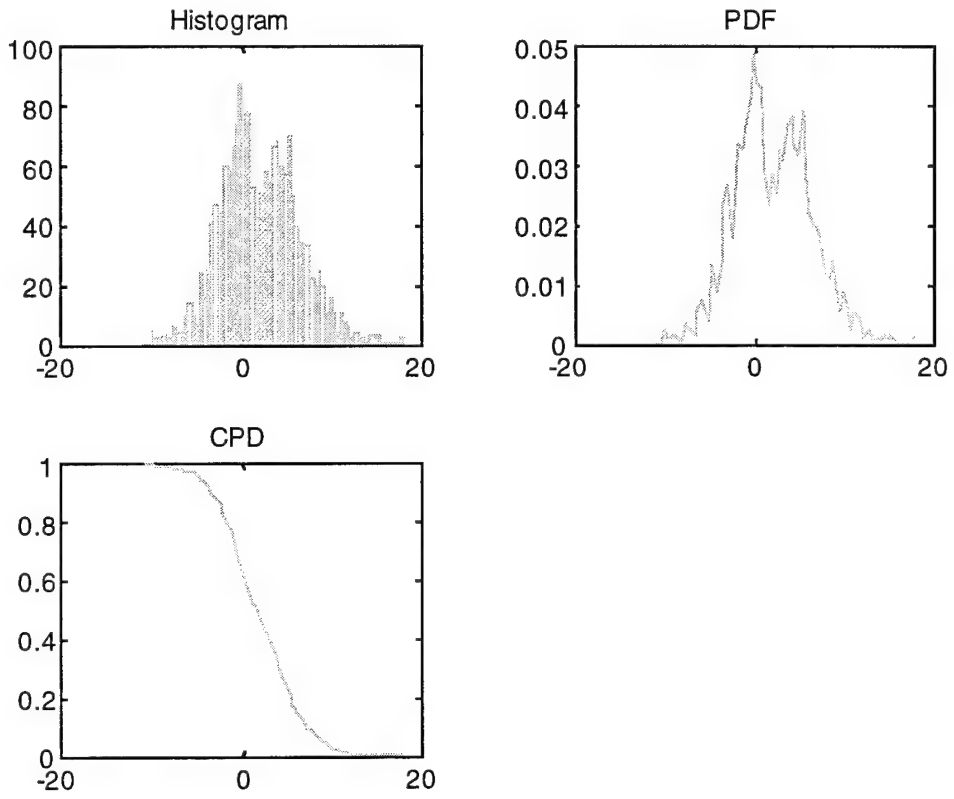


Figure 17. Histogram and Related Plots (Object 1234)

In this case the shape of the histogram is quite different from what was seen for object 2609 (see figure 10). In fact the above histogram appears almost bi-modal. While this may be accidental, it could also indicate the result of receiving returns from two targets within the same range resolution cell of the radar. This problem will be discussed in more detail below. What is significant is the fact that the range of values of the histogram for object 1234 is nearly four times the range seen in the histogram for object 2609. These differences are also seen in the PDF and CPD.

The differences in the RCS time histories are also seen in the plots (figure 18) for the Fourier coefficients.

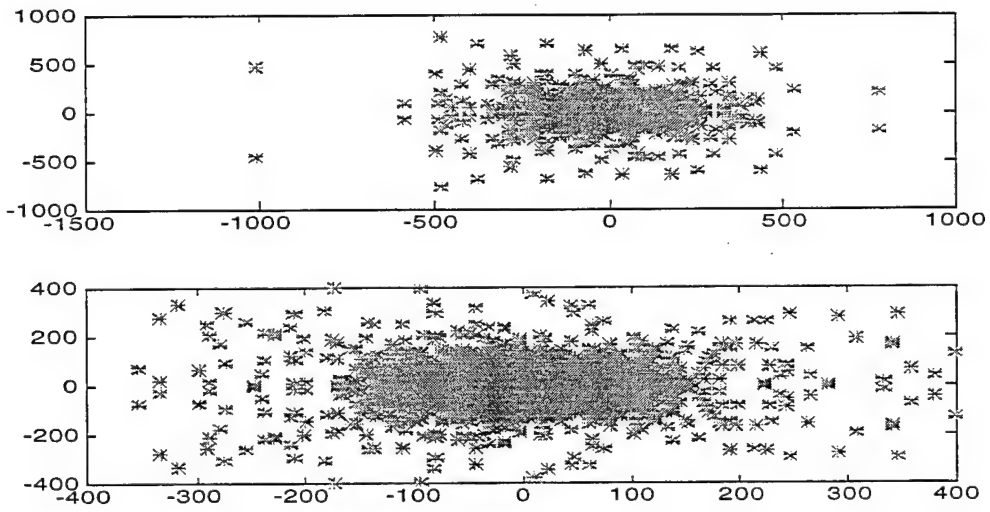


Figure 18. Fourier coefficients (Object 1234)

The spread along both the real and imaginary axis is much more pronounced than for object 2609 (see figure11). Now consider figure 19 which displays the spectral data.

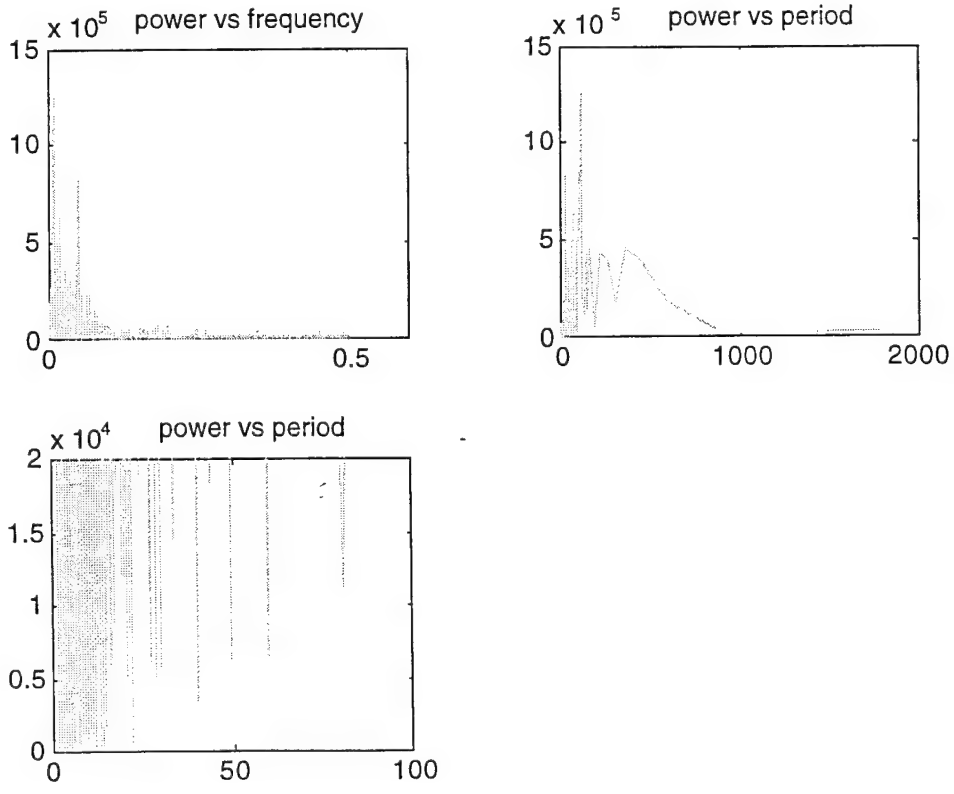


Figure 19. Spectral Data (Object 1234)

There are a couple of things to note. Although it is not obvious due to the way the y axis is scaled, the power in the frequency domain is more than an order of magnitude greater than what was seen for object 2609. Moreover, the linear characteristic in the power vs. period for long periods is absent for this example.

We see that based on a simple classifying scheme, objects with different RCS patterns can both be assigned to the same class. But this is to be expected, since the classifier is sensitive to only the grosser aspects of the pattern, that is the 20 point means and standard deviations. The good news is that we appear to have tools (histograms and spectral representations) which are sensitive to other aspects of the pattern and thus the two data sets are distinguishable. This is not to say that one pattern represents an RV and the other does not. It could well be that both are RVs and the differences in patterns are due to differences in shape, size or body motion relative to the radar line of sight (i.e., tumbling). But we can distinguish the patterns, and when truth data is available, we can begin to assign these patterns to particular objects.

There are of course some issues which need to addressed. First the reader may have noticed that in figures 14 and 15, while the patterns of the RCS were quite similar, the amplitude scales are somewhat different. This could indicate the presence of larger objects that share some of the characteristics associated normally with an RV. The best solution to this difficulty is to establish a threshold based on a reasonable estimate of the RV's RCS. This can be obtained through modeling exercises or from truth data.

A related difficulty is illustrated by figure 20.

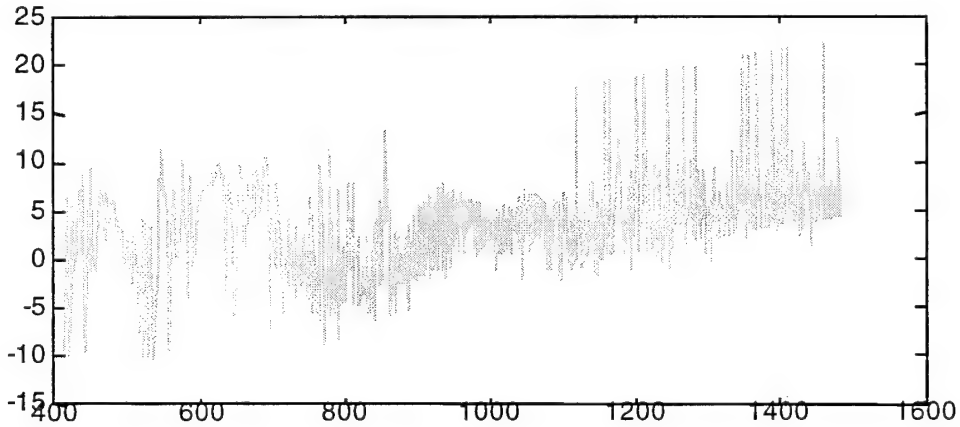


Figure 20. RCS Vs Time (Object 1153)

In this case there is a general increase in the RCS starting at around a time equal to 800 seconds. However, up until that point the pattern is quite different. The problem may be one in which the radar is placing returns from more than one object into the same track file. This problem is termed a misassociation and is not uncommon with these types of radars. However from our point of view, the interesting point is that our routines consider this to be a class 1 object whether there is a misassociation or not. This is the correct assignment, initially at least, since much of the data exhibits a small mean and narrow variance. The fact that the first 400 seconds or so show a different pattern is not important. The fact that there are a small number of high spiky returns at late times is also not important. What is

important is that we do not classify a data set like this as something other than an RV during the first cut.

In any case we should be able to develop algorithms to detect misassociation or at least a discontinuous change in RCS pattern. The late time spikes could be detected by using a high pass filter, but are actually already indicated in the histogram of the data as shown below in figure 21.

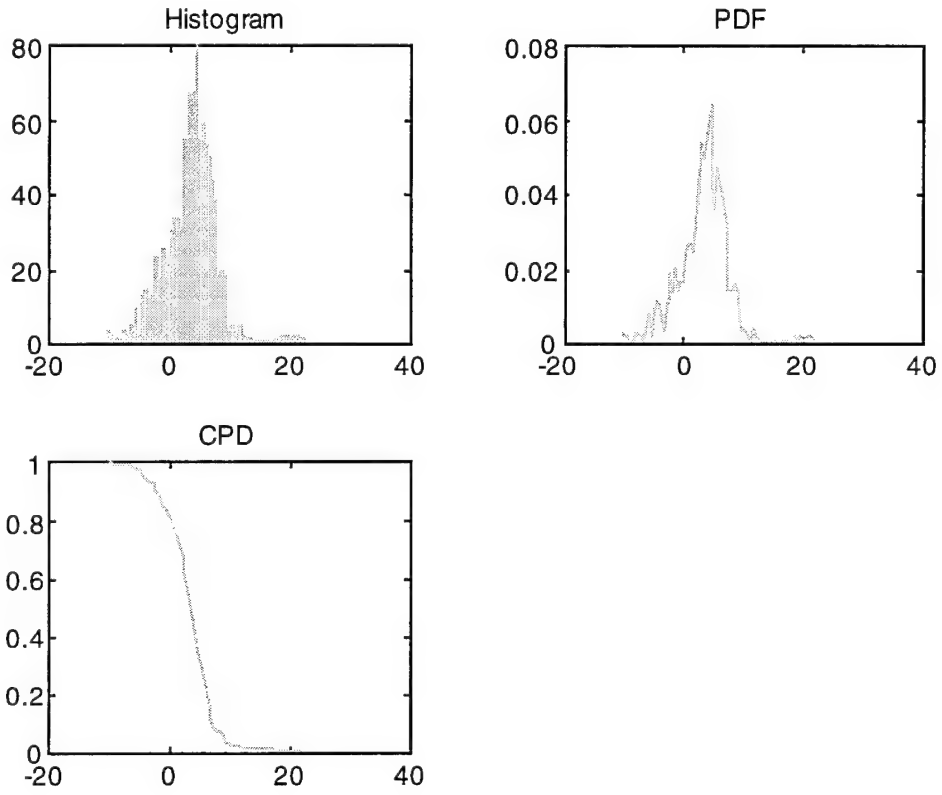


Figure 21. Histogram and Related Plots (Object 1153)

Here the spikes show up in the histogram as a small clustering around bin 20 and are well separated from the main part of the histogram. The spread of the histogram is more than what one might expect for a class 1 or RV object and may be due to the pattern of the RCS returns during the first 400 seconds. In cases such as these, the classifier should probably divide the data set into thirds or quarters and test for a strong 0 Hz spike, which indicates the presence of a systematically increasing amplitude of the return. We simulated this by excising the data from the initial point out to time 800 and then redoing the histogram.

The results are shown below in figure 22.

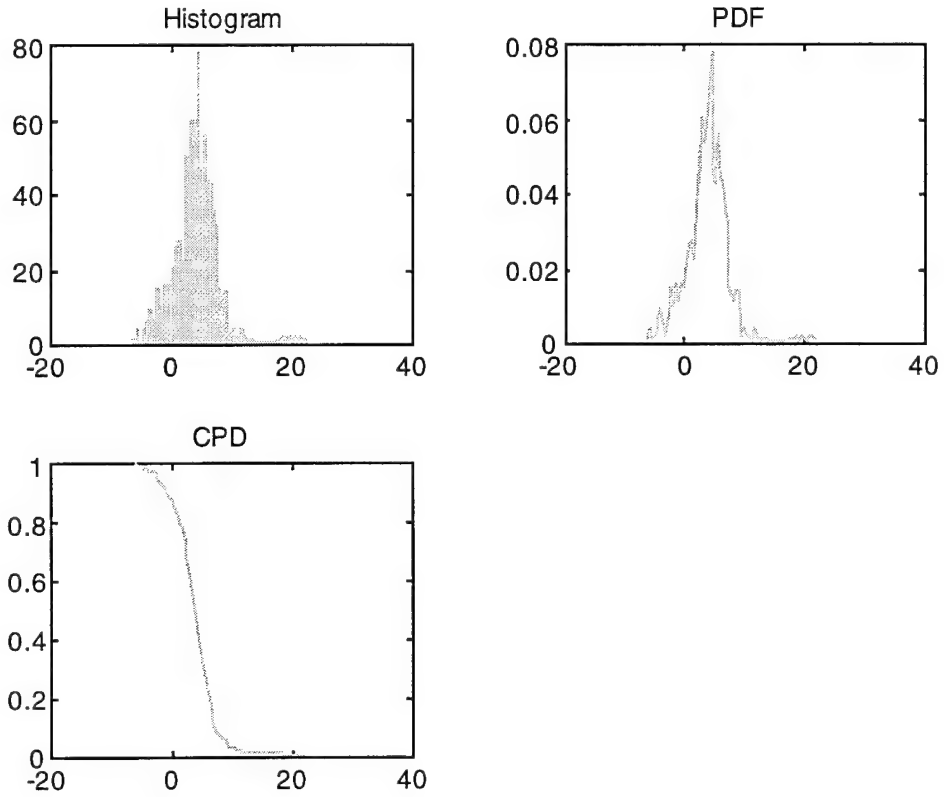


Figure 22. Histogram and Related Plots with the Time Period of 400 to 800 Seconds Removed (Object 1153)

In this case we see a slight difference in the width of the histogram, and we suspect that this data set represents a border line case. In fact our initial classifier is not totally insensitive to this and does give the percentage of subset means and standard deviations which were actually placed in region 1. For the case of object 1153, we noted that only about 50% of these were actually assigned to region 1. Likewise for the data set on object 1234, we discovered that only about 55% were assigned to region 1. On the other hand for the cases of objects 2609, 1193 and 1221, the percentage of mean and standard deviations assigned to class 1 was 88% or higher. So it may well be that our initial classifier can indicate which data sets will require more work. That is which sets may require high pass filtering, searches for 0 Hz lines or discontinuous pattern change identification before a plausible classification can be made.

One final issue must be discussed and this concerns the use of the spectral methods, i.e., the Fourier coefficients, the power vs. frequency plots and so on. The basic process of taking a discrete Fourier transform assumes the time interval between samples is uniform. For a great number of data sets this is simply not the case, even in a rough sense. The following two figures show the RCS vs. time data and the frequency at which the data was collected for objects 1153 and 1229 (figures 23 and 24).

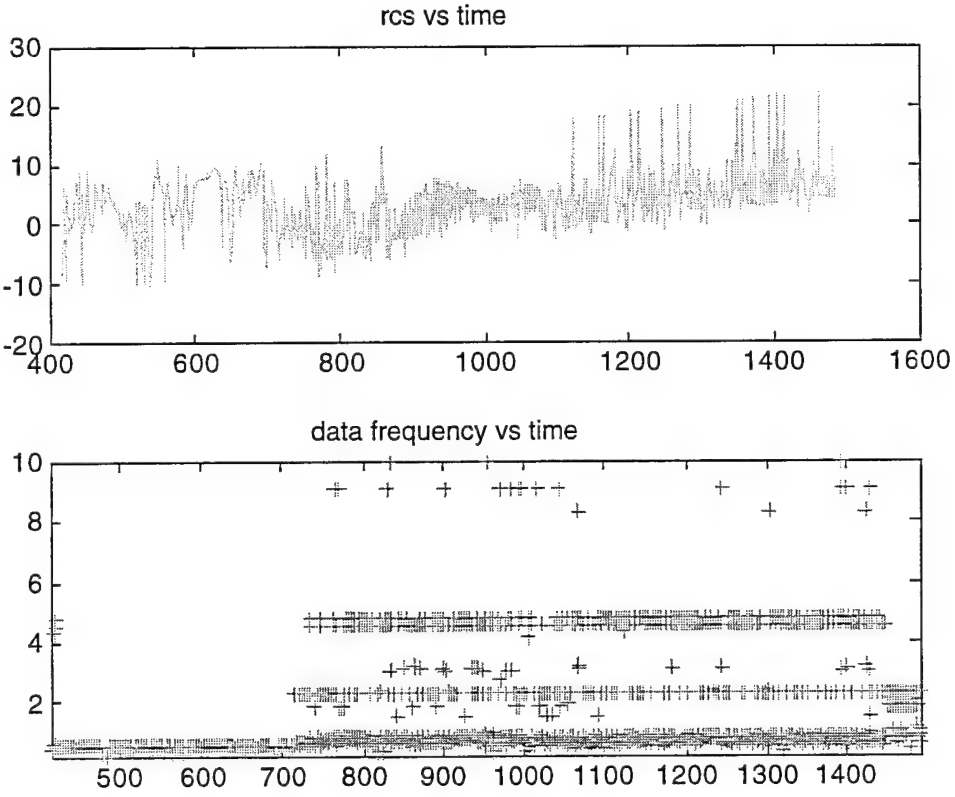


Figure 23. RCS Vs Time and Data Frequency Vs Time (Object 1153)

The first plot gives the RCS time history. The second plot shows the data collection or recording frequency. Thus a value of 10 on the vertical axis indicates a data rate of 10 Hz or a time interval between data points of a 1/10th of a second. At the other extreme, a data frequency of 1/2 Hz means that the time between samples is two seconds. Depending on the physics of the situation, one data rate or another might be preferable. However if the data intervals are constantly changing, then the Fourier transform does not mean very much. Unfortunately, this is precisely what is seen in the above data set - a data interval which changes sporadically with time.

The situation is not completely negative. For example there are some data sets in which the data rates are fairly constant, as shown in figure 24.

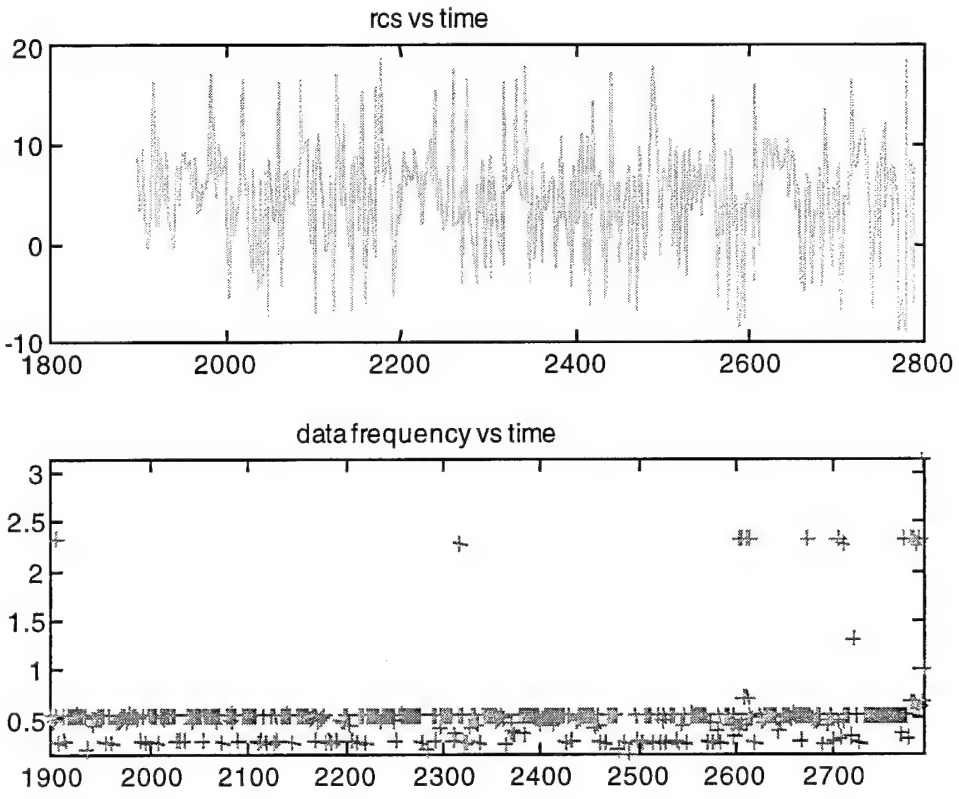


Figure 24. RCS vs Time and Data Frequency vs Time (Object 1229)

It also should be recognized that the information that we are generally seeking from the spectral methods is not of a detailed or quantitative nature. For example, we find that for many of the class 1 objects, most of the power is concentrated at the zero or near zero Hz region. This is a result of a systematic change of the RCS over time and is insensitive to the fluctuating data rate. In addition, we find that for the class 2 or tank like objects, the power tends to be evenly distributed across many frequencies. This is an expected result, because of the essentially random tumbling motion of the tank. It is also a meaningful result because, in general, the data rates for tanks tend to be uniform as in the example shown above.

Discussion

Up to this point, we have considered some initial processing on approximately 52 data sets. This has entailed dividing these sets into four classes based on the 20 point means and standard deviations of their component subsets. This accomplishes two things. First we have the initial cut at the classification in the sense that we now have the data divided into four classes based on some characteristic of their pattern. Secondly it allows us to judge the meaningfulness of our other data processing techniques. That is, do they expose any additional information within the RCS patterns.

At least in the case for class 1 objects, it appears that additional processing serves two purposes. First, the class 1 objects can be divided into further subclasses using data processing techniques that key on other aspects of the RCS pattern. Thus we may have the capability to distinguish between various types of RVs or distinguish RVs from other similarly sized objects.

Secondly we find that by flagging those objects that may be only marginally catalogued as class 1 objects, we can use these other techniques to analyze discrete features of the RCS pattern in an attempt to properly classify the object.

Eventually all of the data will be divided into classes and subclasses, each representing a particular RCS pattern. The real value of this is that the number of distinct patterns appears small and that the patterns repeat themselves over time in the following sense. The data making up groups I and II were collected more than a year apart and yet the same basic RCS patterns exist in both groups. This indicates that if one uses truth data to map the patterns to object types, then this approach has real merit since the patterns are not simply the result of various random processes. This process is schematically summarized in the chart given below.

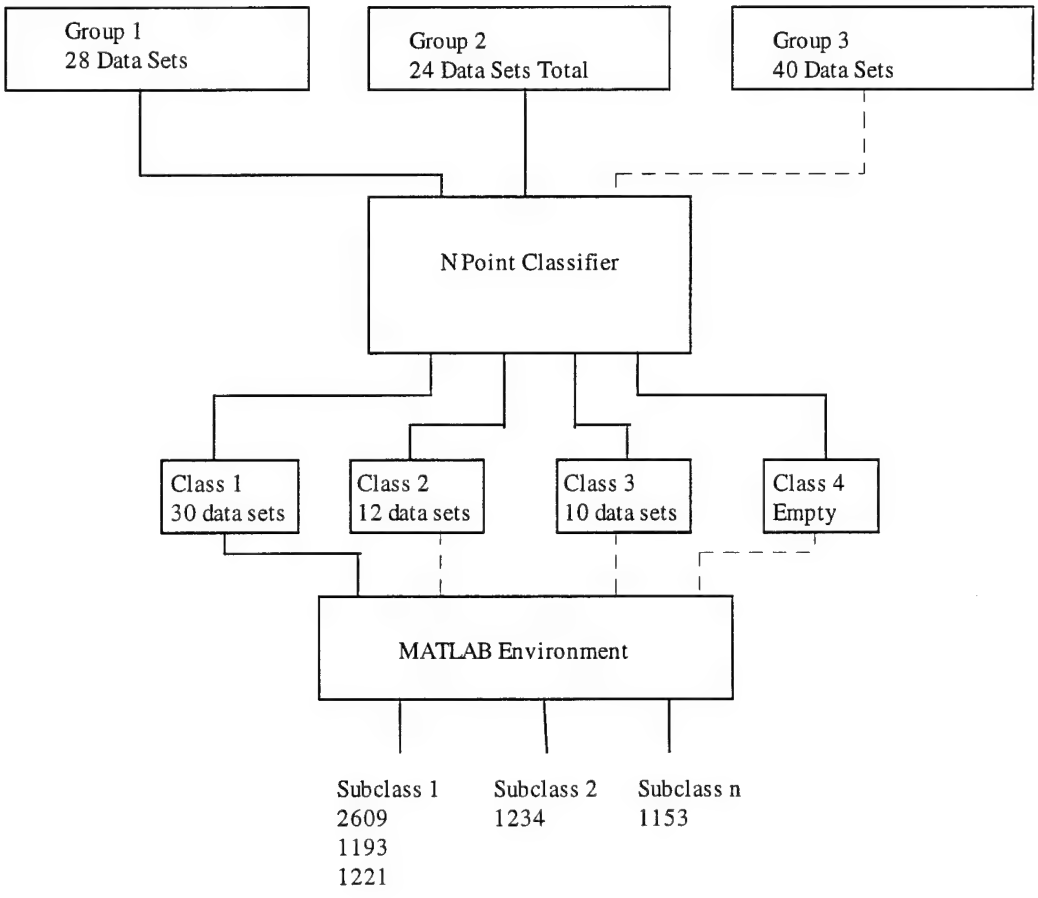


Figure 25. Process Flow Chart

Summary

We can summarize our results as follows. First we divided our RCS data base into three groups, based on collection dates. The data from groups I and II were processed through a primitive classifier which used the means and standard deviations of the component 20 point subsets of each data set. Based on some rough rules of thumb, these data sets were then assigned to one of four classes. Next we incorporated most of the previously developed data processing techniques into a MATLAB frame work. This allows for an efficient and systematic search through the data base. It also represents the first step in establishing a way of emulating the pattern recognition based object classifier.

We have also started to process data from class 1 through the MATLAB environment. This allows us to test the utility of some of our data processing techniques. It has also demonstrated that by using these relatively simple techniques, we can identify other features in the class 1 data for which our primitive classifier was not particularly sensitive. This indicates that once we are able to assign patterns to objects, we may have a methodolgy that can distinguish RVs from similarly sized objects. This hope is reinforced when we note that the patterns are repeated in events that take place years apart.

We have also begun a program to model the RCS returns from certain basic shapes that resemble RVs and tanks (Appendix A). The purpose of this is establish reasonable thresholds on the RCS amplitude and variance. That is, what values of the mean and variance are reasonable when we try to map out the various regions in the mean-standard deviation space of our primitive classifier? We also expect that this approach can shed some insight in to how the RCS pattern changes when one allows the RV to execute certain types of tumbling motion.

Future Efforts

Our next task will be to finish processing the data through the MATLAB environment. This will further demonstrate the utility of our techniques and more importantly allow us to identify and characterize all of the significant patterns within our data base. This will be the key result, since it indicates that only a finite set of patterns exist and that simple methods or techniques can identify them. Once this is done, the identification or pattern recognition process will be formalized into the algorithmic structure of a classifier. At this point we hope to have some truth data to calibrate our patterns to actual object types.

In addition to this main line of development, we also intend to evaluate the usefulness of the higher order moments of the data sets (i.e., skewness and kurtosis) for our pattern recognition process. We also intend to continue our RCS modeling work. This will include setting up the problem of determining from the RCS data, the appropriate aspect angle history, and finding the target which has the pattern that most likely would give the data observed. We do this by calculating the gradient of the RCS pattern with respect to aspect angle times the gradient of aspect angle with respect to motion parameters This is simply the gradient of the RCS with respect to motion parameters which can be used in the gradient search algorithm to find the motion parameter vector that makes the RCS model as close as possible to the measured data in the least squares sense.

Appendix A

RCS Simulation

In this appendix we discuss our initial attempts to generate simulated RCS data under a wide range of target and trajectory parameters. This should give us some feeling for what the RCS patterns may look like under a broad range of cases, in particular that of a tumbling RV. Another reason for considering this particular simulation is that it leads to an interesting pattern recognition scheme which is discussed below. The ideas presented here are somewhat speculative, however their potential payoff makes them worth considering.

Method of Moments Electromagnetic Scattering Model

A computer program developed by Mitzchang and Putman at McDonnell-Douglas employing the method of moments (MOM) has been obtained and setup for use in this study. The program calculates the RCS of a user specified body of revolution (BOR) composed of conductors, dielectrics and cavities (as appropriate). Since RVs and tanks can be modeled as BORs, this program is quite relevant. Moreover, MOM gives a solution to the scattering problem that accurately satisfies the boundary conditions on the electromagnetic field imposed by the body composition, so that a nearly exact representation of the RCS is available from the program. Calculations involving MOM are feasible only when the body is no more than a modest number of wave lengths in size. As it turns out, the parameters of the radar and target are such that MOM is an ideal analysis tool.

The MOM solution is found by assuming that the incident electromagnetic field has illuminated the body and has induced a surface current distribution which then must be calculated. This current distribution re-radiates a "scattered" electromagnetic field that is received by the radar and from which the RCS is calculated. The total electromagnetic field (the sum of the incident and scattered fields) must satisfy both Maxwell's equations and all boundary conditions imposed by the environment, which includes the scattering body and free space. Thus a consequence of applying Maxwell's equations in integral form to a region including a boundary between a dielectric and a perfect conductor, is that the tangential electric field is zero on the boundary and the tangential magnetic field is discontinuous by an amount proportional to the induced current.

The re-radiated field can be expressed as a linear superposition (an integral) of the fields radiated by basic current elements making up the induced current distribution on the conductor boundary. Both the electric and magnetic fields can be expressed by such an integral. An integral equation for the current is obtained by evaluating the fields on the boundary and imposing the boundary conditions for that field. If the integral expression for the electric field is used, the boundary condition sets the tangential component of the electric field to zero on the boundary. The resulting integral equation for the current is called the Electric Field Integral Equation (EFIE). When the integral expression for the magnetic field is used, the boundary condition sets the tangential magnetic field equal to the induced boundary current. In this case the resulting integral equation for the current is called the Magnetic Field Integral Equation (MFIE). The solution of either of these equations gives the current on the boundary, although the accuracy and stability of each solution can vary and depends on the configuration of the scattering problem.

The method of solving either one is essentially to convert the integral equation into a matrix equation, which may be solved to evaluate the coefficients of a series expansion of the unknown current. The current is then evaluated by summing the series expansion for the current using the newly found coefficients. Once the currents are known so is the scattered field and the RCS.

A method combining the two equations called the Combined Field Integral Equation (CFIE) turns out to be the most flexible method to solve these types of problems. This approach was implemented by Mitzchang and Potter and resulted in a program called CICEROC.FOR. The results of using this program are discussed below.

Possible Classification Scheme

CICEROC produces an output of RCS values as function of aspect angle. Typically the values chosen by a user would be from 0 to 180 degrees, in equal angular increments. However, in general a radar will sample RCS values of a target for only a finite set of aspect angles as determined by the orientation of the target body relative to the radar line of sight, the motion of the target relative to the radar, and the sampling interval. Thus a typical RCS history obtained by a radar would only include RCS samples from a subset of the aspect angle set that the analytical model (CICEROC) can produce, since this program produces RCS values for all aspect angles desired.

Now suppose we were to represent the motion of the body using a set or vector of parameters that described the body motion and the actual history of the orientation of the body. Then any set of motion parameters or more properly any vector of motion parameters would have a one to one correspondence with a history of aspect angles that the target presents to the interrogating radar. Since we do not know the aspect angle history of the actual data, we can not match it to a library of RCS vs. aspect angle values that can be obtained from CICEROC. However, we can develop a program that would vary the motion parameter vector (and hence the corresponding aspect angle history) so that the difference between the actual measured RCS data sequence and one obtained from the library corresponding to the aspect angle history is minimized. This quantity is then minimized over the set of all motion parameter vectors, for each target model in the library (tank, RV, etc.). The smaller the difference, the more likely the data came from a particular target model being tested, with the specified motion vector or aspect angle history.

Thus we have a two step process. First by minimizing over motion parameter space, for a fixed test target, we get the most likely body dynamics given that that target is actually the one observed. Next we minimize over targets and we find the most likely target that fits the data.

Some Preliminary Results

We have generated simulated RCS data that would be observed by a radar illuminating a moving target. Two target models were created, including a tumbling and spin stabilized RV and one tank type. The RV was modeled as a parabolic cylinder with a conducting tail piece, represented by a disk at the rear of the RV. To simulate a tank section, we chose a simple straight cylinder with nose and end (disk) caps. Other RV shapes will be modeled later.

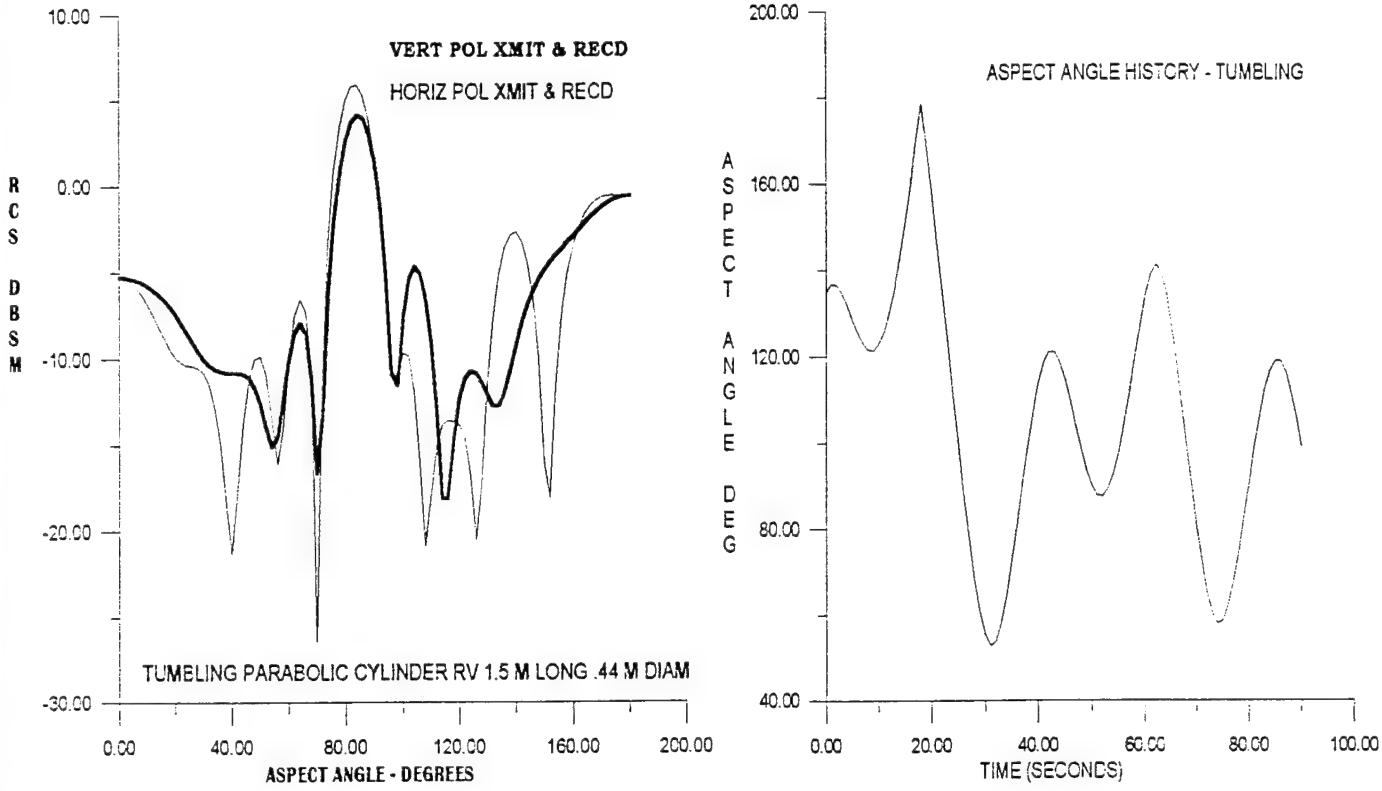
A total of 12 parameters were used to delineate a target motion vector. They are: $\vec{R}_0$, the initial target location relative to the radar, the target velocity, $\vec{v}$, the target acceleration, $\vec{a}$

and roll, pitch and yaw angles, (R, P and Y) as linear functions of time. A program called MOMACH.FOR calculates the aspect angle relative to the radar due to a prescribed motion. The motion is prescribed by setting the 12 parameters. The motions available with such a parameter vector space will be ballistic transport, accompanied by roll, pitch and yaw with constant angular change rates. Most RVs are spin stabilized or are tumbling in some prescibed fashion. The motion parameter space as described will be detailed enough to simulate such motions.

One of the classification methods that is being examined under this study is the use of cumulative probability distribution functions for the RCS data treated as a set of statistical samples of a random process. A histogram of the data set is made giving the relative frequency of occurrence of a particular value of RCS in the data sample. This is an approximation to the probability density function for RCS in the data. We then integrate this function up to a particular RCS in the dynamic range of the data to get the probability that the RCS is less than that value, and subtracting from unity gives the probability that the RCS at an sample will be greater than that particular value. This is called the cumulative probability distribution (CPD) and is calculated for each data set. As mentioned before, the CPDs for tanks and RV's have distinct features that can distinguish them. For example RV's are usually spin stabilized and smaller in RCS than tanks, which are not spin stabilized and tend to rotate or tumble as they progress along their ballistic trajectories. Thus the mean RCS for an RV is usually less than that of a tank and the variance of the tank is usually greater than that of the RV because of the entire range of RCS values can be sensed from a tumbling object, whereas a small range of RCS values may be acquired from a spin stabilized object.

However, the mean and variance provide only a limited characterization of the behavior of a random process that is non Gaussian. The actual shape of the CPD can often distinguish targets, because they are derived from the actual RCS values for the targets which will be quite different for different targets. This is because the currents induced on the targets (and the corresponding resulting fields) are quite different and depend on the target shape and composition, at least for targets that are not too small electrically. As the wavelength goes up, and the target becomes electrically smaller, we expect the scattered fields for the different targets to become more similar. However, for the radar under investigation, and the targets of interest, RV 's are often slightly larger than the wavelength and tanks are much larger. This means that we are in the resonance region of scattering where changes in target shape can alter the scattered field markedly.

In Figure 26a, we show 3 plots pertaining to the monostatic RCS of a tumbling object resembling an RV. It was modeled as a parabolic cylinder with a length of 1.5 meters and a tail diameter of .444 meters. The illuminating frequency is 430 MHz, corresponding to a wavelength of .6977 meter. Thus the RV model is 2.15 wavelengths long and .636 wavelengths in diameter at the tail.



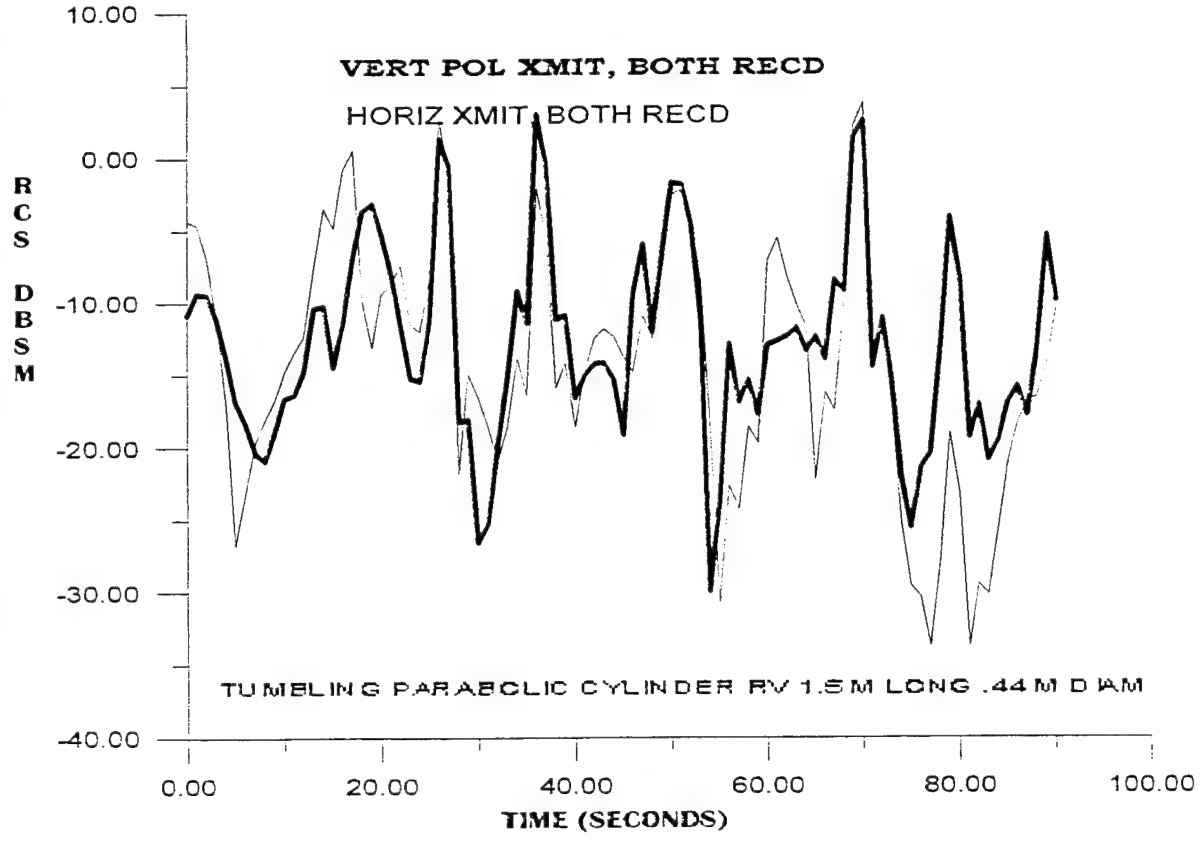


FIGURE 26a RCS vs aspect, aspect history, and RCS vs time for tumbling RV

The RCS as a function of aspect angle was computed from CICEROC and is shown in the upper left corner of igure 26a. The RCS is given for both horizontal (H) (light curve) and vertical (V) (dark curve) polarizations, with the body axis parallel to the horizontal direction. The aspect angle is defined as the angle included between the radar line of sight (RLOS) and the axis of symmetry of the BOR, with 0 degrees aspect when the RLOS is parallel to the axis of symmetry and pointed toward the nose of the body (the RV pointing towards the radar). The plot exhibits a large lobe on broadside incidence (90 degrees aspect), a slightly smaller but still large lobe in the region of tail incidence (180 degrees aspect) and a smaller and wider lobe at the nose.

The second plot, in the lower half of the figure, shows the RCS as a function of time when the body moves and tumbles. The sampling interval in all cases was one second. The initial range was taken at 20 kilometers, with an initial azimuth angle at -60 degrees, and an initial elevation of 45 degrees. The velocity of the target was taken to be 20 meters/sec in the x-direction (transverse to radar line of sight) with a 20 m/s/s acceleration in the same direction. The body was assumed to rotate at a uniform rate through 720 degrees of pitch angle and 720 degrees of yaw angle, with no roll angle. These motion parameters lead to the aspect angle history shown in figure 26a at the top right corner. For this motion, almost the entire body is viewable by the radar at some time.

Figure 26b displays the CPD for 4 cases. The plot in the upper half of the figure is the CPD for horizontal polarization, while the vertical polarization is shown in the lower half. In both plots, the dark curves are CPD obtained using full aspect angle dependent RCS data as presented in figure 26a in the top left corner. This would represent the CPD for data obtained if the RCS for each aspect angle was claculated by the radar. Thus the dark curve CPD curves represent the CPD given that any aspect angle was equally likely to be observed by the radar. The lighter curves represent the CPD obtained from the radar data when the object moves and tumbles according to the aspect angle history given in figure 26a (upper right corner plot). The CPD so obtained is conditioned on the distribution of aspect angles that the motion presents to the radar, and as the aspect angle history shows not all aspect angles are equally likely.

Note that for this case, the square root of the sum of the squares of the horizontal and vertical received signals are plotted, regardles of which polarization was transmitted. This is because, since the body tumbles, the horizontal polarization from the ground becomes a mixed polarization when the body is illuminated, since the body orientation is in general variable along its trajectory due to its motion relative to the transmitted field. Thus two polarizations are reflected and when received by the radar, we assume that the radar is equipped with receivers at both horizontal and vertical polarizations. To insure maximum RCS visibility, we assume that the receiver simply evaluates the magnitude of the received field vector by combining its polarizations using the Pythagorean theorem.

Figures 27a and b show the same set of plots as given in Figures 26a and b, for an RV that is spin stabilized. For this case, the roll, pitch and yaw angles of the body are fixed throughout the motion. One notices immediately that the dynamic range of RCS values sampled in the time data plotted in the lower half of figure 27a have been markedly reduced from the tumbling RV case. This is due to a the restricted span of aspect angles viewed by the radar. Of course, the dynamic range could be larger if the range of aspect angles included a region with both a lobe peak and a null. The CPD plot (horizontal polarization) shown in the upper half of figure 27b is similar to that of the tumbling RV except that the curves have a much steeper slope. Again this is due to the limited range of aspect angles viewed in the data. The vertical polarization case in the lower half of figure 27b shows a behavior similar to the tumbling RV case in figure 26b because the vertical polarization

RCS samples as shown in figure 27a (lower half page) span almost the entire range of RCS values for the body. This is because the aspect angle history for the spin stabilized case includes a region where the RCS goes from a peak to a null, for vertical polarization.

Figures 28a and b show the set of plots for a tumbling right circular cylinder, representing a tank section. Here the cylinder is 2 meters in diameter and 6 meters long. The dynamic range or variance of the RCS in both the full pattern and the time sampled data is much larger than that for either the tumbling or spin stabilized RV, and its mean is much larger. Thus the CPD's for the tank are wider in span and centered about a larger RCS mean value. The tumbling tank is easily distinguishable from either the tumbling or spin stabilized RV due to these differences.

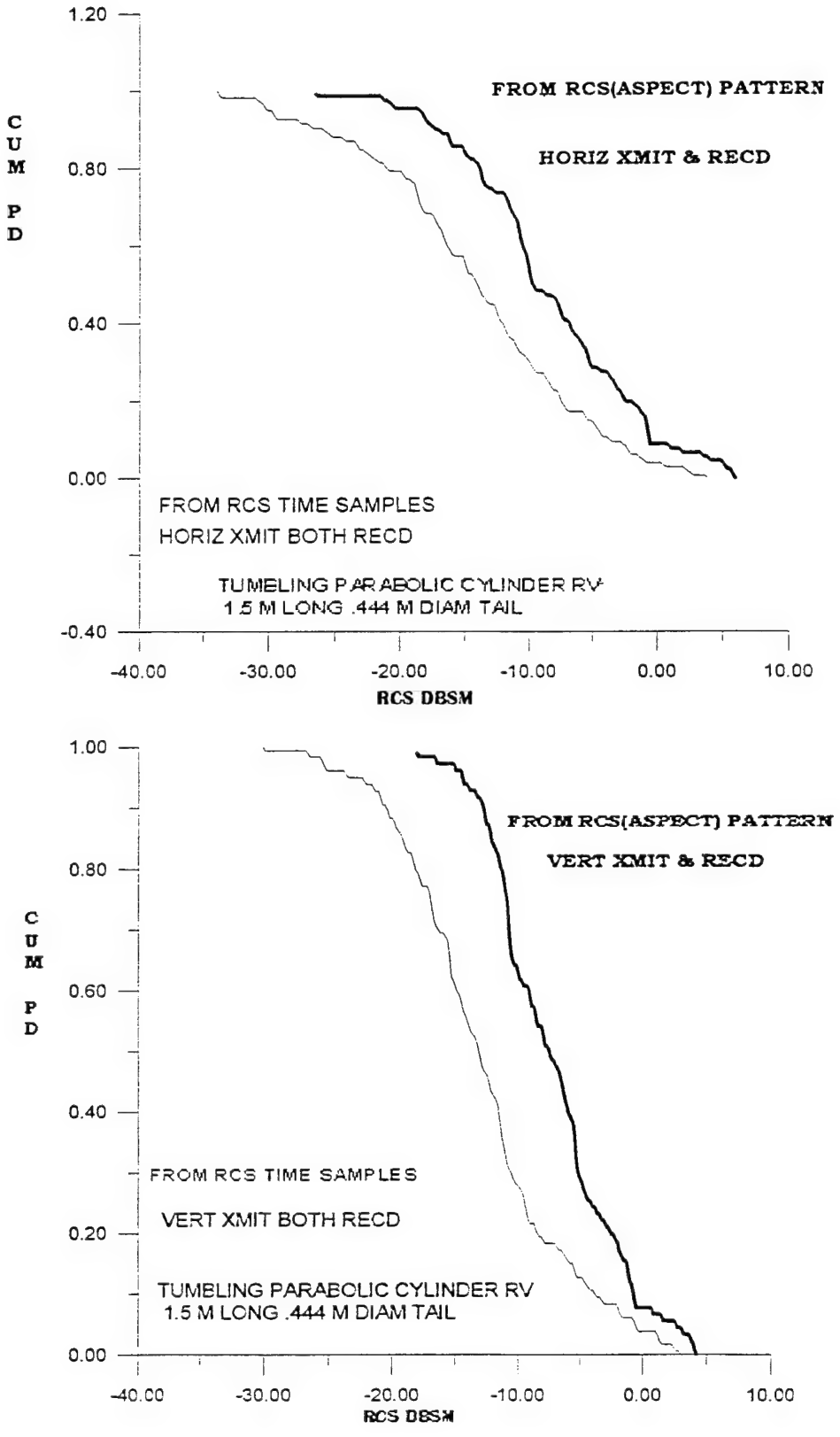


Figure 26b Cumulative probability distributions.
Comparing pattern generated and time sample generated.

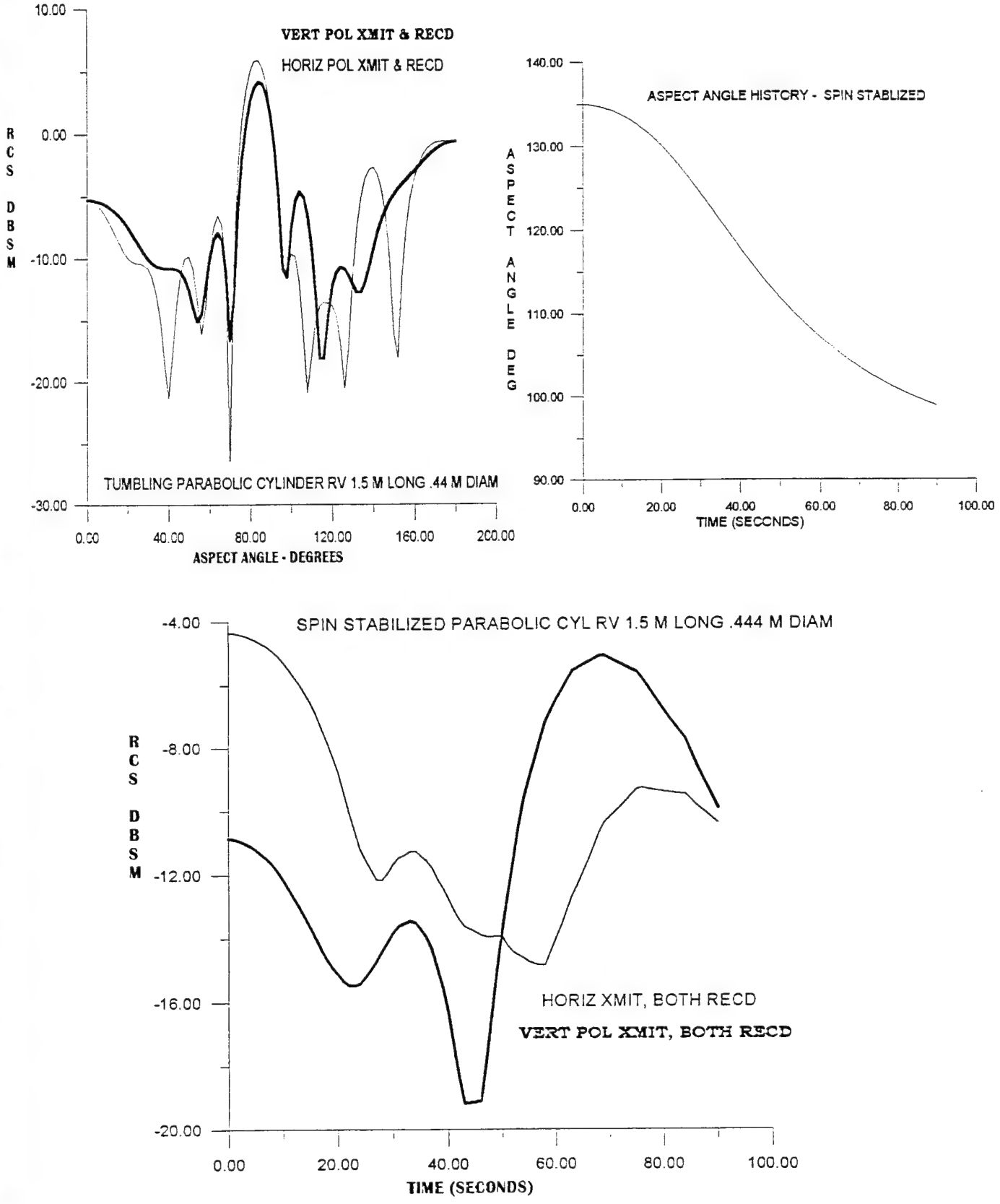
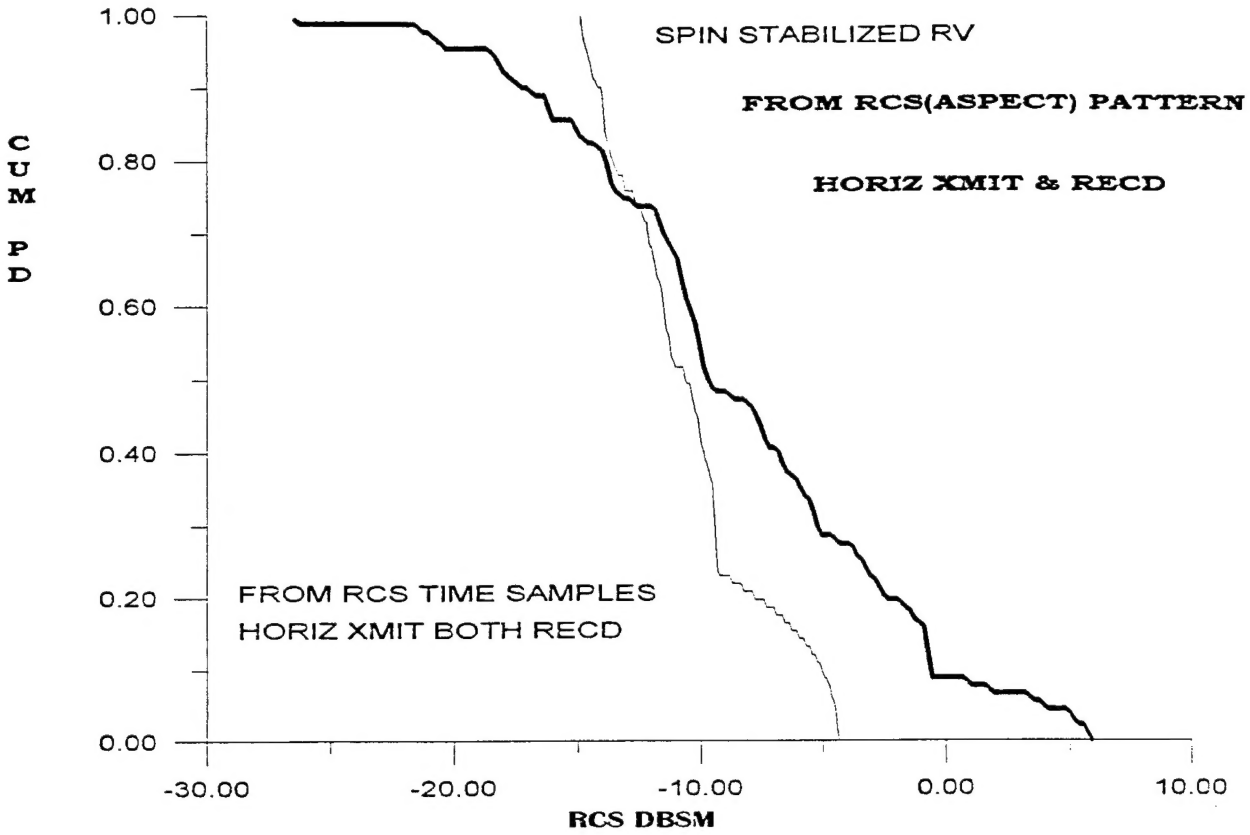


FIGURE 27a RCS vs aspect, aspect history, and RCS vs time for spin stablized RV



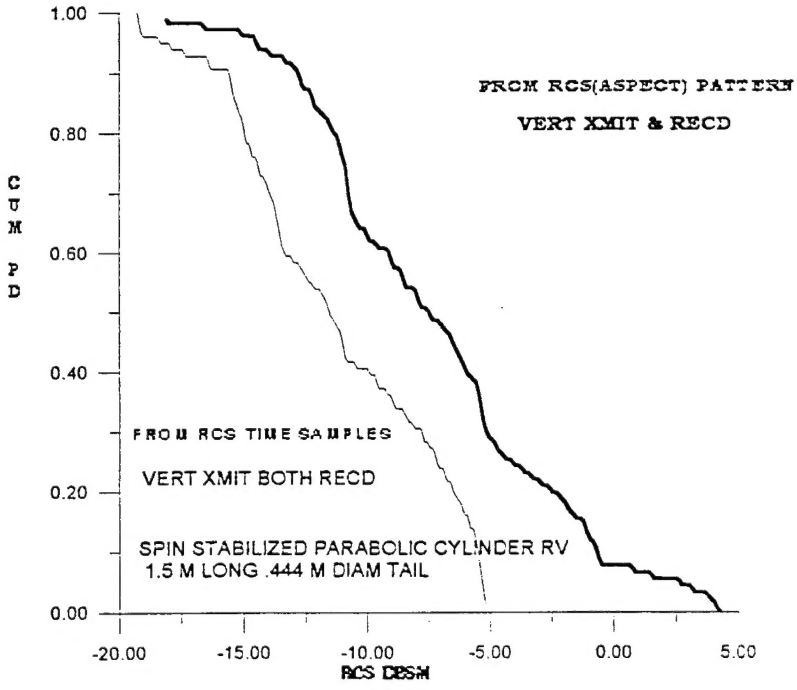


Figure 27b Cumulative probability distributions.
Comparing pattern generated and time sample generated.

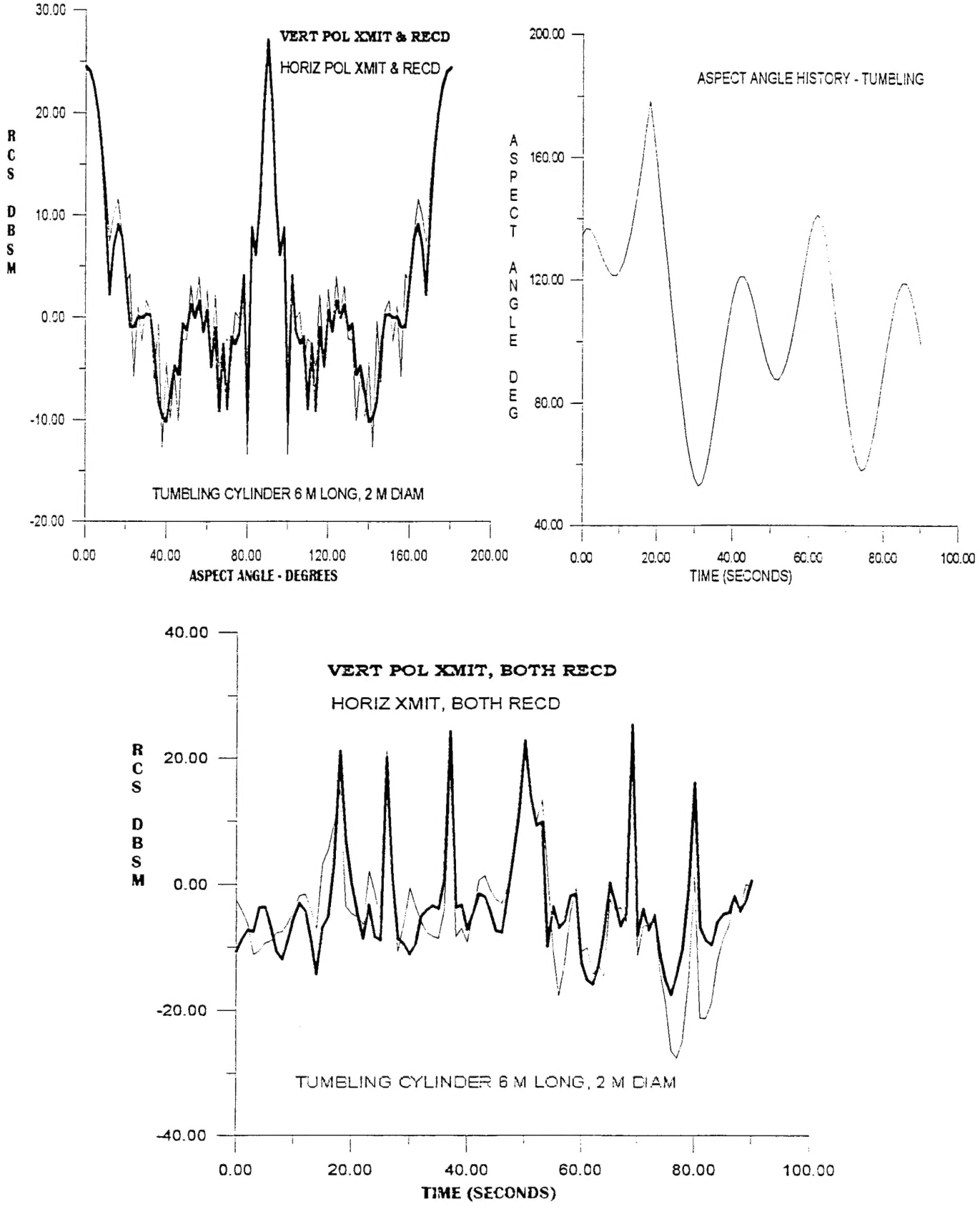
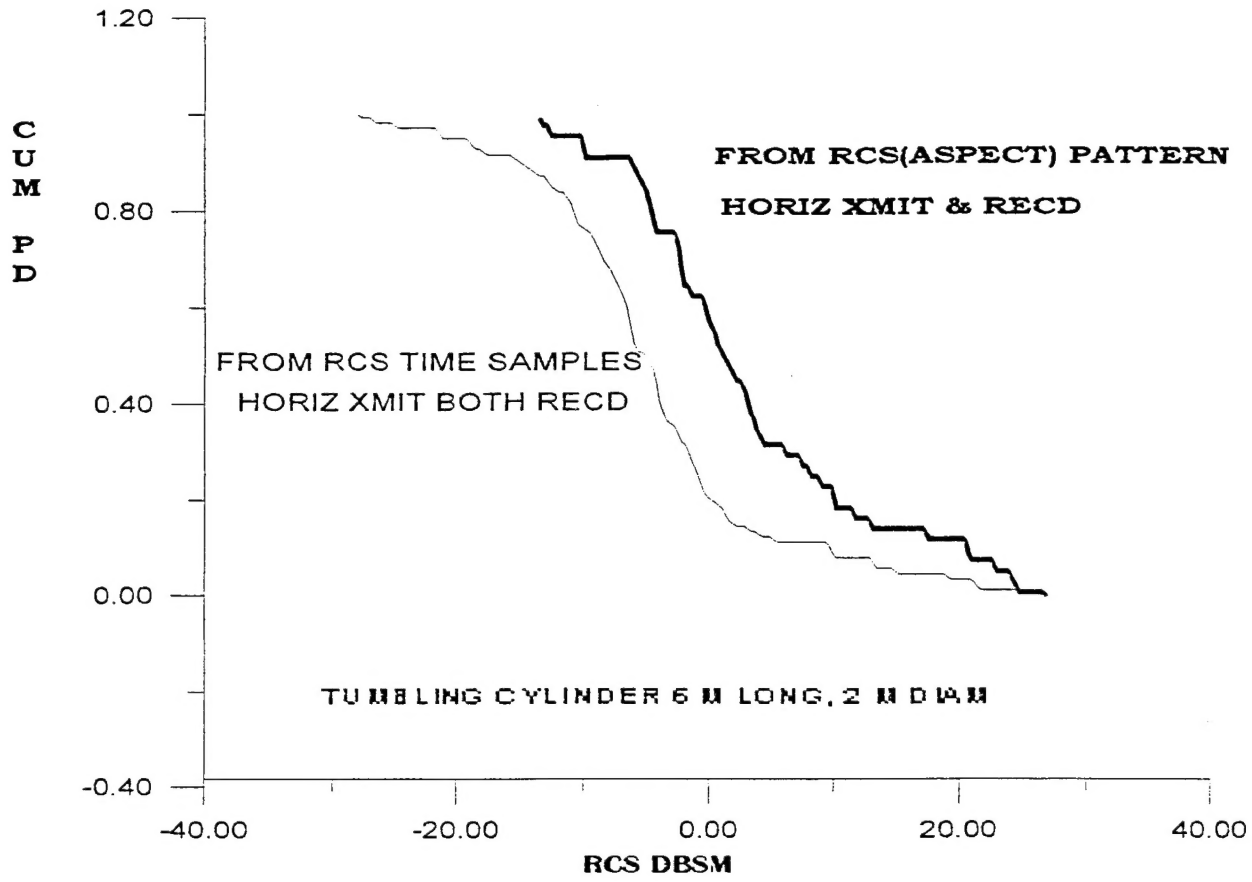


FIGURE 28a RCS vs aspect, aspect history, and RCS vs time for tumbling cylinder



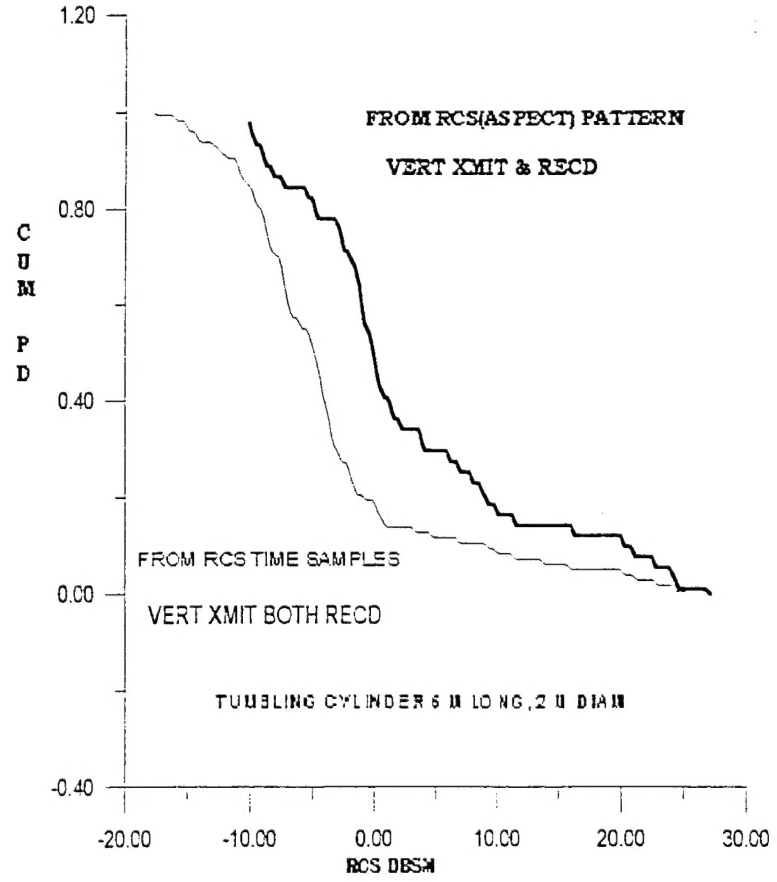


Figure 28b Cumulative probability distributions.
Comparing pattern generated and time sample generated.

Discussion

We now present a few qualitative remarks to put the above discussion in perspective and summarize our objectives in terms of this RCS modeling effort. Aside from the possibility of using this RCS prediction routine in a pattern recognition scheme, there are two reasons for pursuing this approach. First we would like to have a way to establish an RCS threshold in terms of the mean and variance. For example, we find patterns in the RCS data base that closely resemble that of an RV, but its overall RCS amplitude is higher than what we might expect. In the absence of truth data, we might be able to use the RCS modeling tool to establish upper limits on the RCS amplitudes relative to an RV.

A second reason is to consider the problem of a tumbling RV. We would like to use this tool to establish what we might expect the RCS signature to look like in terms of the RCS amplitude and variance.

We have not yet accomplished either of these two goals, although we have made progress. From the point of view of establishing a threshold, it appears that all of the RV RCS amplitude results are scaled down from what we think we see in the in the actual data. The source of this discrepancy is not easily determined. However, an obvious place to consider is the modeling of the RV in terms of its size and shape. It may well be that we need to model this much more accurately than we done so far.

On the other hand, we see qualitative agreement between the CPDs generated by the RCS model and those produced from the actual data. For both the tank and RV cases, the slopes and shapes of the CPDs are quite close to what we have seen in the data. We also see qualitative agreement in the RCS amplitude of the tank targets between what is seen in the data and what is predicted by the model.

In terms of establishing the characteristics of the tumbling RV, we probably face the same problem of a lack of fidelity in modeling the shape and size of the RV. However, the peak RCS is certainly enhanced for the tumbling case over the spin stabilized situation. By comparing the RCS vs time plots in figures 26a and 27a, we see that the peak RCS for the tumbling case (figure 26a) is approximately 8 dB above what I seen for the spin stabilized case show in figure 27a.